**Title**
Essays on Online Job Search



**Author**
Zhang, Shuo

**Publication Date**
2022

Peer reviewed|Thesis/dissertation

University of California
Santa Barbara

# Essays on Online Job Search

A dissertation submitted in partial satisfaction
of the requirements for the degree

Doctor of Philosophy
in
Economics

by

Shuo Zhang

Committee in charge:

Professor Peter Kuhn, Chair
Professor Kelly Bedard
Professor Clément de Chaisemartin

June 2022

The Dissertation of Shuo Zhang is approved.

Professor Kelly Bedard

Professor Clément de Chaisemartin

Professor Peter Kuhn, Committee Chair

June 2022

Essays on Online Job Search

To my parents, Qinghe and Lihong.

## Acknowledgements

There are many people who helped me along the way on this memorable journey. First and foremost, I am extremely grateful to my advisor, Prof. Peter Kuhn, who has gone over my manuscripts word for word, and corrected them from a comma to the whole perspective, from my second-year proposal to my dissertation, who encouraged and recommended me to attend conferences and workshops, and who is always willing and enthusiastic to provide any helps he can in my study and life. His immense knowledge, plentiful experience, and unsurpassed insights guide me to find the right path to a research question in the shortest time, and motivate me to fully exploit the potentials of my research. He is my role model as a scholar, a teacher, and an advisor, whom I will consistently learn from in my future work and life.

I would like to express my heartful gratitude to Clément de Chaisemartin for entrusting and offering me the RAship on our R package projects. I am deeply grateful for his solid and insightful comments on my research, and his continuous and unqualified support in my job market. I also appreciate all the encouragement and inspiration that I received from Kelly Bedard, and her remarkable contributions to our department as a Chair. She taught me how to balance career and family, and always guided me in a positive way. I would like to thank the faculty members at UCSB, especially Erik Eyster, Heather Royer, Shelly Lundberg, and Catherine Weinberger for their kindness and feedback.

In addition, I would like to offer my special thanks to my cohort and department colleagues, Lizi Yu, Yizhou Liu, Di Wang, Zheng Wei, Guangli Zhang, and Emily Robertson for their friendship and company; they are always willing to listen to me and always be there when I need a hand. I had the pleasure of collaborating with my coauthors, Kailing Shen, Donald Lien, and Taoxiong Liu, and I would not be the researcher I am today without their help and patience. I also wish to acknowledge my colleagues, Yi Shan, Hua Zhao, and Xueyan Xu at Liepin.com, for providing me with invaluable research data and generous assistance as much as they can from their busy schedules.

And my biggest thank goes out to my family. During the years in pursuit of my degree, I have missed many events, and my parents, Qinghe and Lihong shouldered the entire burden of the family. Your unconditional love and endless support endow me with the courage to go further, and I am forever grateful for your sacrifice and understanding. I would like to thank my best half, Sicong, for steering me through the difficult times, and constantly supporting my career financially, technologically, and emotionally without reservation. Meeting you is the best thing that ever happened to me at UCSB. Last but not least, I want to thank my cutest daughter, Grace, for choosing me as your mom and giving me unlimited happiness. Raising you reminds me of how I grew up, and teaches me how to be a good person. I love you with everything I am.

**Curriculum Vitæ**
Shuo Zhang

**Education**

| | |
|---|---|
| 2022 | Ph.D. in Economics, University of California, Santa Barbara. |
| 2020 | M.A. in Statistics, University of California, Santa Barbara. |
| 2016 | M.A. in Economics, University of California, Santa Barbara. |
| 2015 | B.A. in Economics, Law, Tsinghua University. |

**Research Fields**
Labor Economics, Behavioral Economics, Applied Econometrics

**Publications**
"Gender-targeted Job Ads in the Recruitment Process: Facts from a Chinese Job Board", with Peter Kuhn and Kailing Shen, *Journal of Development Economics 147 (2020): 102531*.
"Words Matter Life: The Effect of Language on Suicide Behavior", with Donald Lien, *Journal of Behavioral and Experimental Economics 86 (2020): 101536*.

**Working Papers**
"Measuring Algorithmic Bias in Job Recommender Systems: An Audit Study Approach".
"Should I Show or Should I Hide–When Do Jobseekers Reveal Their Wages", with Peter Kuhn, Taoxiong Liu and Kebin Dai.

**Other Projects**
R Package Development: $DIDmultiplegt$, $TwoWayFEWeights$

**Teaching Experience**
Teaching Assistant, University of California, Santa Barbara

| | |
|---|---|
| Intermediate Microeconomic Theory | *2020* |
| Introduction to Econometrics | *2019* |
| Personnel Economics | *2018* |
| Financial Management | *2017* |
| Principles of Economics-Microeconomics | *2016* |
| Principles of Economics-Macroeconomics | *2015* |

**Honors and Fellowships**

| | |
|---|---|
| Summer Fellowship, UCSB | *2021* |
| UCOP Diversity RAship, UCSB | *2020* |
| Departmental Research Funding, UCSB | *2020* |
| White Family Graduate Fellowship, UCSB | *2019* |
| First-Class Comprehensive Scholarship, THU | *2014* |

## Permissions and Attributions

The content of Chapter 1 and Appendix A is the result of a collaboration with Peter Kuhn and Kailing Shen, and has previously appeared in the *Journal of Development Economics*. The journal allows the inclusion of previously published work in a student's dissertation with proper attribution. The published article is available at https://doi.org/10.1016/j.jdeveco.2020.102531
Peter Kuhn, Kailing Shen, Shuo Zhang, "Gender-targeted job ads in the recruitment process: Facts from a Chinese job board", Journal of Development Economics, Volume 147, 2020, 102531, ISSN 0304-3878.

The content of Chapter 2 and Appendix B is the result of a collaboration with Peter Kuhn, Taoxiong Liu and Kebin Dai.

**Abstract**

Essays on Online Job Search

by

Shuo Zhang

This dissertation focuses on the job searching and matching in online labor markets. Based on the field experimental data and internal data from job boards, my dissertation explores research topics including workers' job search behaviors, employers' recruitment decisions, and the role of internet job platforms in online job matching.

In the first chapter, joint with Peter Kuhn and Kailing Shen, we study how explicit employer requests for applicants of a particular gender enter the recruitment process on a Chinese job board, focusing on two questions: First, to what extent do employers' requests affect the gender mix of a firm's applicant pool? Second, how 'hard' are employers' stated gender requests– are they essential requirements, soft preferences, or something in between? Using internal data from a Chinese job board, we estimate that an explicit request for men raises men's share in the applicant pool by 14.6 percentage points, or 26.4%; requests for women raises the female applicant share by 24.6 percentage points, or 55.0%. Men (women) who apply to gender-mismatched jobs also experience a substantial call-back penalty of 24 (43) percent. Thus, explicit gender requests do shape applicant pools, and signal a substantial but not absolute preference for the requested gender.

The second chapter, based on joint work with Peter Kuhn, Taoxiong Liu and Kebin Dai, studies how workers make voluntary wage disclosure decisions in the job search process using internal data from a leading online Chinese job board, Liepin.com. We find that on average, workers' disclosure decisions are consistent with a model in which

high current wages are seen as "good news" by prospective employers: Workers are more likely to disclose their wages when their wages are higher than might be expected, based on the worker's resume and where they applied. Employers' responses to workers' resumes, however, are hard to reconcile with these disclosure patterns: While employers respond positively to workers with higher-than-predicted current wages, they do so equally, *regardless of whether those wages have been disclosed*. This suggests that firms can infer the unobserved ability associated with a worker's current wages from other aspects of her resume and application behavior. Finally, the act of disclosing one's current wage –regardless of its level– appears to *reduce* firm's interests in hiring a worker. Disclosures of low wages (which are rare) appear to be mistakes (because they reduce both application success rates and offered wages); disclosures of high wages may, however, benefit workers by filtering out unwanted low-wage job offers.

The third chapter investigates gender bias in job recommender systems. By conducting an algorithm audit in four Chinese job boards, I find that gender-specific jobs, which are only displayed to one gender, account for 9.72% of the total recommended jobs to identical male and female applicants. Gender-specific jobs differ in both the job's explicit quality and the words used in job descriptions: Compared to jobs that are only recommended to men, only-to-women jobs propose lower wages, request fewer years of working experience, are more likely to require literacy skills and administrative skills, and tend to contain words related to feminine personality, which reflect gender stereotypes in the workplace. Item-based collaborative filtering, content-based recommendation algorithms and the hiring agents' behaviors incorporated in job recommender systems are the possible drivers of the gender bias in job recommendations.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Gender-Targeted Job Ads in the Recruitment Process: Facts from a Chinese Job Board

## 1.1 Introduction

Statements in a job ad that either men or women are preferred by the employer are used in many developing-economy labor markets; these include India, Indonesia, Brazil, Pakistan, Nigeria, Russia, Mexico, Colombia, Peru and others which account for substantial shares of the world's labor force.[1] Recent studies of posted job ads, including Kuhn and Shen (2013); Helleseter et al. (2020); and Ningrum et al. (2020) have demonstrated the following about these ads: Gendered job ads are much more com-

[1] Appendix A.1 provides examples of explicitly gendered job ads from the ten most populous countries served by Indeed.com ("the world's 1 job site"), representing 57 percent of the world's population. With the exception of the United States, gendered ads were easy to find on all the remaining platforms. A similar search on Computrabajo.com (which serves 20 Spanish-speaking countries) quickly detected explicit gender requests on all the larger platforms – including Colombia, Mexico, Argentina, Peru and Venezuela – with the exception of Spain and Chile.

mon in jobs requiring low versus high levels of skill, and the gender requested in a job ad is more closely tied to the job's duties than to the identity of the firm posting the ad. Gendered ads tend to reinforce, rather than counteract existing stereotypes of male and female work, with requests for men most common in jobs like construction, driving, security services and upper management, and requests for women dominating in 'helping' and customer contact jobs like receptionists, clerks and customer service. Finally, employers' explicit gender requests are highly correlated with requests for other employee attributes: Jobs that request women are much more likely to request younger, attractive, single workers, while ads that request men tend to request older, married workers.

While the above research provides new information about how *when and how often* employers post gendered job ads, to our knowledge no research has yet studied how these ads enter the recruitment process after they are posted. In particular, economists still lack answers to two key questions: First, how do workers respond to gender-targeted job ads? Do these ads direct workers' search toward jobs that request the worker's gender, and away from jobs that request the opposite gender? This question is interesting because it sheds light on *why employers use* gendered job ads. Second, how 'serious' are employers when they make a gender request in a job ad? At one extreme, advertised gender requests could be hard requirements in the sense that gender-mismatched applications are always rejected, or are successful only when no workers of the requested gender apply. At the other extreme, advertised gender requests could just be soft suggestions that a particular gender is preferred, or even that a particular gender might prefer working in that job (for example due to the presence of same-sex co-workers or a flexible work schedule). This question is interesting because it sheds light on *why workers comply* with employers' gender requests, and on the extent to which gendered job ads limit men's and women's choices in the labor market.

To address these questions, this paper uses internal data from a Chinese job board (XMRC.com) to establish a set of basic facts about how explicit gender requests in job ads enter the recruitment process. A key advantage of our data is that – in addition to knowing the characteristics of all the ads (including the requested gender, if any) – we know the gender and qualifications of every person who applied to each ad, and (for a subset of the ads) the gender and qualifications of the persons who were called back to the ad. We establish four facts about aggregate patterns, and document two partial correlations that suggest causal effects of explicit gender requests on application and callback behavior.

First, as a summary indicator of the extent to which employers' personnel selection decisions reflect their gender requests in the job ad, we ask the following question: If a job ad requests a particular gender, what share of its successful applicants (in our case, callbacks) are of that gender? This statistic, which we refer to as *gender matching* – is 94 percent in jobs requesting women, 96 percent in jobs requesting men, and 95 percent overall. Thus, 19 in 20 callbacks to gendered job ads are of the requested gender. Second, a key source of this high gender matching rate is self-selection by workers: 92.5 percent of *applications* to gendered job ads are of the requested gender; this number – which we refer to as workers' *compliance* with employers' gender requests – is very similar for jobs requesting men versus women. Notably, these matching and compliance statistics for employers' gender requests are higher than the corresponding statistics for employers' age, education and experience requests, suggesting that employer's gender requests play a particularly important role in the matching process.[2]

Third, both men and women who apply to jobs that request the opposite gender experience lower callback rates than workers who apply to non-gendered jobs, or to

[2]See Section 2 for our exact definitions of matching on these dimensions. For example, in the case of age our preferred measure is the share of callbacks that fall into the age range that is explicitly requested in the job ad.

jobs requesting their own gender. Thus, at least in the aggregate statistics, employers appear to *enforce* their own gender requests by penalizing gender-mismatched applicants. This enforcement is far from absolute, however. For example, among applicants to jobs requesting women, men are 80 percent as likely to get a callback as women. Among applicants to jobs requesting men, women are only 45 percent as likely to be called back as men, a difference which is highly statistically significant. Thus, at least in the aggregate statistics, women who apply to 'men's' jobs succeed much less frequently than men applying to 'women's' jobs.

Fourth, decomposing the total amount of gender matching in the aggregate data into components associated with compliance, enforcement and their interaction, we find that these components account for 74, 6, and 20 percent respectively. Intuitively, the dominant role of compliance reflects the fact that applicant pools to explicitly male and female jobs are highly gender-segregated. Thus, if these application patterns are (hypothetically) held fixed, the gender mix of callback pools would strongly match employers' requests even if hiring from applicant pools was gender neutral in all job types.[3]

Fifth – and turning now our motivating questions – , the high level of workers' apparent compliance in the aggregate statistics is not just an artifact of the tendency for, say, women to apply to stereotypically female jobs (which request women more frequently in our data). To demonstrate this, we regress the female share of applications to a job ad on indicators for whether the ad requests men or women, with controls that include firm-by-job-title fixed effects. Thus, even when comparing job ads posted by the same firm for the same job title, we estimate that adding an explicit request for men to a job ad reduces the female share of applicants by 15 percentage points, or 26 per-

[3] We emphasize the descriptive nature of this decomposition because high self-sorting could be caused by high enforcement.

cent; a request for women raises the female applicant share by 25 percentage points, or 55 percent.[4] Importantly, Marinescu and Rathelot (2018*a*) show that job titles are more detailed and more predictive of wages and application decisions than are six-digit SOC codes.

To shed additional light on how employers' explicit gender requests interact with job titles in influencing workers' application decisions, we use a Bayesian machine learning approach (McCallum et al., 1998) to identify job ads whose gender preferences can be clearly predicted from the job title, and those that cannot. Consistent with the hypothesis that prospective applicants try to infer their hiring prospects from all the information contained in the ad, we find that explicit gender labels have the largest effects on applicant gender mix in jobs whose title does not suggest a clear gender preference on the employer's part.[5] Further, we find that men and women respond differently to this ambiguity: essentially, men are not deterred from applying to 'gender-ambiguous' jobs, while women tend to apply only when their gender is explicitly requested. This pattern – which echoes existing findings that female job searchers are more ambiguity-averse, and more responsive to affirmative action statements than men (Gee, 2019; Ibañez and Riener, 2018) – accounts for the larger effect of female than male labels on the gender mix of applicants.

Finally, we show that the substantial apparent enforcement by employers of their own gender requests in the aggregate statistics is not an artifact of how workers of different ability levels self-select into making gender-mismatched applications. To demonstrate this, we regress an indicator of whether an application received a callback on indicators for the six possible matches between worker types (men and women) and

[4]Consistent with Kuhn and Shen (2013) model of the effects of advertised employer preferences, requesting either male or female applicants has a cost on XMRC: it reduces the total number of applications received. Effects of gender requests on the observed quality and match of applications (on dimensions other than gender) are robustly zero, however.

[5]Some common job titles with this feature are "international trade person" and "accountant".

job types (male, female, and no gender request), with fixed effects for job titles and for individual workers. Also included are detailed controls for firm and job characteristics, and for the match between the job's requirements and the worker's qualifications. Thus, even when comparing applications made by the same worker to the same job title, to which the worker is identically matched according to education, experience, and age requirements, we estimate that gender-mismatched applications experience a substantial callback penalty.[6]

Specifically, we estimate that a man's callback probability falls by 2.2 percentage points (or 24 percent) if he applies to an identical, explicitly female job compared to a nongendered job. Women's callback chances fall by a greater amount (3.7 percentage points or 43 percent) when applying to an explicitly male job compared to a nongendered job. While highly statistically significant, both these effects are smaller in magnitude than the corresponding regression-unadjusted differentials, a fact that sheds light on the nature of selection into gender-mismatched applications. For example, women who apply to jobs requesting men might do so primarily because they feel they are better qualified according to some other characteristic – such as education or experience – that compensates for being of the 'wrong' gender. If so, selection into gender-mismatched jobs would be positive, and controlling for resume fixed effects would increase the size of the estimated mismatch penalty. Instead, we find that the estimated penalty falls, implying negative selection. This suggests that workers who apply to gender mismatched jobs are of lower ability, or apply for jobs more indiscriminately than other workers.

Our paper contributes to a number of literatures, the first of which uses the contents of job ads to study labor markets. These studies include Hershbein and Kahn (2018)

[6] We also control for the relationship between the applicant's current (or most recent) wage and the wage advertised in the job ad.

and Modestino et al. (2016), both of which ask whether employers request higher qualifications for the same jobs when local labor market conditions make workers "easier to get". **?**Brenčič (2010); Brenčič and Norris (2012), and Brenčič (2012) use the same type of data to study aspects of employers' recruiting strategies, including whether to post a wage and whether to adjust ad contents during the course of recruitment. More recently, analysis of posted job ads has played a rapidly growing role in the analysis of developing-country labor markets, partly in response to gaps and weaknesses in government-run surveys of employment and vacancies. These include Nomura et al. (2017) and Ahmed (2018) for India; Matsuda et al. (2019) for Pakistan; and Hayashi and Matsuda (2020) for Bangladesh and Sri Lanka, all of whom use job board data to study detailed and high-frequency changes in employment and skill demand. Relative to all these articles, a key advance of our paper is the use of internal job board data to see whether and how such changes in ad content actually matter: do they direct workers' search, and do they inform potential applicants of how employers will respond when workers who do not meet the advertised criteria apply?

Second, our paper relates to a large literature that studies racial, gender, and other differentials in callback rates using resume audit methods (Bertrand and Mullainathan, 2004; Kroft et al., 2013; Neumark et al., 2019). While our estimates of callback differentials are not experimentally based, a key advantage of our job-board-based approach is that it lets us study callbacks to the entire population of jobs on offer, which vary dramatically in their gender preferences. For example, even though a roughly equal number of jobs on XMRC request women and men, 85 percent of ads for front desk personnel explicitly request women, and 88 percent of ads for security personnel explicitly request men (Helleseter et al., 2020). This extreme heterogeneity poses a challenge for audit studies, which typically elicit an average race or gender preference in a relatively narrow set of jobs, often selected to be approximately race- or gender-neutral.[7] In contrast, a key parameter in our approach is this heterogeneity, as captured by our *mismatch penalty* parameter: how does, say, a woman's callback probability change when she redirects her application from a nongendered to an equivalent female job? As already noted, our estimates of the mismatch penalty control for unobserved worker quality by using worker fixed effects, since we can observe the same worker applying to different types of jobs.

Another related literature is a rapidly growing group of empirical papers that study where jobseekers decide to send their applications. Motivated in part by an older theoretical literature on directed search in labor markets (e.g. Albrecht et al. (2006)), these papers include Marinescu and Wolthoff (2020); Belot et al. (2017); and Banfi and Villena-Roldan (2019), all of whom study the effects of the posted wage on the number and quality of applications a firm receives. Marinescu and Rathelot (2018*a*) study the geographic scope of workers' search, and Kudlyak et al. (2013) study how workers re-direct their search over the course of a search spell. Ibañez and Riener (2018) and Leibbrandt and List (2018) study the effects of affirmative action statements on application decisions, while Flory et al. (2015) and Mas and Pallais (2017) study how workers' application decisions respond to competitive work environments and non-wage job attributes respectively.[8] Our paper differs from these in at least two key ways: it is the first to focus on the effects of explicit gender requests in ads, and – instead of focusing on a very particular subset of jobs – it studies application and callback decisions in the

[7]In addition to cost, a key reason for this narrow focus is the difficulty of constructing plausible resumes for a large variety of jobs, many of which are highly specialized. Thus, for example, both Bertrand and Mullainathan (2004) and Kroft et al. (2013) restrict their attention to four occupations: sales, administrative support, clerical, and customer service. Carlsson and Rooth (2007) study is noteworthy for studying the heterogeneity in discrimination across 13 occupations.

[8]An emerging concern in this regard derives from the increasing capacity to micro-target all types of online ads. For example, Verizon recently placed a job ad that was set to run "on the Facebook feeds of users 25 to 36 years old who lived in the nation's capital, or had recently visited there, and had demonstrated an interest in finance"(Angwin et al., 2017). In contrast to the Chinese case that we study – where all applicants can view all ads – in the Facebook case non-targeted workers were not even aware of the ad's existence.

entire population of ads on this job board.

Finally, there is a large literature on gender differentials in labor markets, but very little of it has focused on the explicit gender profiling of jobs in emerging economy labor markets like the one we study here. Understanding this practice would seem to be an essential component of understanding gender differentials in labor markets in much of the world. We hope that this paper, which establishes a first set of basic facts about how gendered ads enter the recruitment process in these markets, will stimulate additional research on this under-researched phenomenon.

Section 1 of the paper describes our data source. Section 2 presents aggregate estimates of gender matching, compliance and enforcement. Sections 3 and 4 conduct regression analyses of compliance and enforcement respectively. Section 5 discusses implications of the results and possible avenues for further research.

## 1.2 Data

As noted, our data consist of internal records of XMRC.com, an Internet job board serving the city of Xiamen. XMRC is a private firm, commissioned by the local government to serve private-sector employers seeking relatively skilled workers.[9] Its job board has a traditional structure, with posted ads and resumes, on-line job applications and a facility for employers to contact workers via the site. XMRC went online in early 2000; it is nationally recognized as dominant in Xiamen, possibly due to its close links with the local government and social security bureau.[10]

[9]The other major local job site, XMZYJS, is operated directly by the local government. It serves private sector firms seeking production and low-level service workers. Unlike XMRC, XMZYJS does not host resumes or provide a service for workers and firms to contact each other through the site.

[10]XMRCs offices are in the same building as complementary local government offices (e.g. for social security and payroll taxation), offering employers the advantage of 'one-stop shopping' for employment-related services.

To document how gendered job ads enter the recruiting process on XMRC, we began with the universe of ads that received their first application between May 1 and October 30, 2010. We then matched those ads to all the resumes that applied to them, creating a complete set of applications. Finally, for the subset of ads that used XMRC's internal messaging system to contact applicants, we have indicators for which applicants were contacted after the application was submitted. This indicator serves as our measure of callbacks. Our primary dataset for the paper is this subset of ads where both application and callback information is available (henceforth the callback sample), which comprises 3,637/42,744 = 8.5 percent of all ads. In Section 3, however – where we focus only on application behavior—we use the full sample of 42,744 job ads. Appendix A.4 provides summary statistics for the full and the callback samples; they are very similar. Appendix A.4 also replicates our analysis of application behavior on the smaller, callback sample, with very similar results.

Aside from being the only integrated dataset of ads, resumes, applications and callbacks we are aware of – especially in an environment that permits gendered job ads – , an important advantage of our 2010 XMRC sample is its simple and unambiguous indicator of employers' gender requests. On many job boards (both in China and elsewhere), employers' gender requests must be inferred by parsing the text of the ad, a process which requires a number of judgment calls.[11] On XMRC, in contrast, when creating a profile for each new job that is advertised, employers were given the option to specify a desired gender. This datum was then displayed in the job's online description, together with (and in the same format as) more standard desiderata like education and experience requirements, which are collected in the same way. Thus, our measure of whether the employer states a gender preference is simple and stan-

[11] For example, in Spanish one must decide whether "abogada" and "abogado" as job titles are explicit gender requests; in Chinese one must decide whether the adjective 'beautiful' can describe both men and women.

dardized across all job ads.

A second advantage of our setting is the relatively simple nature of the search technology on the site: In 2010, XMRC's site largely emulated printed job ads, where workers peruse ads using simple search filters to decide where to apply. More recently (and coming soon to XMRC), many job boards use machine learning to display suggested job matches to individual workers based on the worker's location, qualifications, employment history and recent searches. In these cases, the jobs a worker applies to are jointly determined by the jobs that are suggested to her by the board's algorithms and her choices from that set.[12] This joint determination does not apply to our data.

Third, the environment in Xiamen in 2010 was remarkably free of legal impediments to posting a gendered job ad, and free of stigma attached to employers posting such ads. While China's constitution has formally given women equal rights since 1982, these principles had few practical consequences for labor markets until July 2012, when the first lawsuit claiming gender discrimination in employment was filed. The first regulations that appear to have constrained firms' ability to post gendered job ads on online job boards appeared in May 2016, when China's Ministry of Industry and Information Technology clearly specified fines for both job boards and employers posting such ads.[13] Since then, some Chinese job boards (especially some prominent national boards) responded by eliminating – or at least making it hard to find – overtly discriminatory job ads on their sites. Smaller and regional job boards continued to post explicit gender requests after 2016, but enforcement has been increasing; XMRC finally removed explicit gender requests in March 2019.[14] That said, as described in

[12] We do not observe which ads were viewed by workers; thus our estimated effects should be interpreted as incorporating workers' decisions regarding which types of jobs to search for. See Horton (2017) for a recent analysis of the effects of algorithmic recommendations in the labor market.

[13] See Appendix A.2 for additional details on China's labor laws as they apply to gender profiling in job ads.

[14] See Appendix A.3 for a recent survey of gender targeting on Chinese job boards.

Appendix A.3, even boards that have eliminated gendered ads continue to allow indirect signals of their employers' desired gender, such as "gentleman" (绅士), "beautiful face" (面容姣好), and "little brother/sister" (小哥哥) which refers to attractive young men and women. Perhaps more importantly, many sites still allow recruiters to filter applications and resumes by gender, making it easy to restrict their attention to only male or female applicants.

In sum, while gendered recruitment by employers is still present in China's new legal environment, it is now less overt, more varied in form and harder to detect. XMRC in 2010 thus provides a picture of how employers would choose to advertise jobs when unconstrained, and of how employers treat applications that do not match a measure of gender preferences that employers have few incentives to misrepresent. Arguably, our XMRC data may also provide insights for how gendered job ads work in countries where they remain largely unregulated.

In all, our primary dataset comprises 229,616 applications made by 79,697 workers (resumes) to 3,637 ads, placed by 1,614 firms, resulting in 19,245 callbacks. Thus there was an average of 63 applications per ad and 5.3 callbacks per ad. One in twelve applications received a callback, while one in four resumes received a callback. Notable features of the job ads (documented in Appendix A.4) include the fact that 867/3,637 = 24 percent of ads requested female applicants, 18 percent requested male applicants and the remaining 58 percent did not specify a preferred gender. The average years of requested education were 12.2, and were more than a year higher in jobs requesting women than men. Forty-eight percent of ads specified a preferred worker age; the mean requested age was 28. Consistent with the age twist identified in Helleseter et al. (2020), the requested age was considerably lower for jobs specifically requesting women. On average, one year of experience was requested. 58 percent of ads posted a wage; the mean posted wage was 2,446 RMB per month overall but only 2,001 RMB in jobs requesting women.[15]

Notable features of the applications (also from Appendix A.4) are that 124,275/229,616 = 54 percent of applications came from women. The typical application had 14.35 years of education, with women holding about half a year more education than men. Average applicant age was 24.0 years. Other applicant characteristics observed in our data (and used in the regression analysis) include experience, new graduate status, marital status, current wage (when provided), myopia, height, the number of experience and job spells listed, and whether an English version of the resume is available.

To provide some context for the sample of jobs and workers on XMRC, Appendix A.4 also compares the characteristics of job ads on XMRC with those of private-sector employees in Xiamen and in urban China.[16] These samples differ quite dramatically: the ads on XMRC seek workers who are considerably younger, better educated, better paid, and more female than the employed population of Xiamen, or of a typical large Chinese city. This is as we might expect, for three reasons. The first is XMRC's explicit niche in the local labor market: to serve relatively skilled workers. Second, due to a massive recent expansion of China's higher education system, highly skilled workers tend to be very young.[17] Third, as on any job board, the ads and resumes on XMRC represent vacancies and jobseekers, not employed workers. Thus we would expect new labor market entrants (who are all looking for work) and young workers (who turn over more frequently than other workers) to be substantially overrepresented relative

[15] Non-gendered jobs pay more than both F and M jobs because they have considerably higher skill requirements. Thus, the comparison between F and M jobs of 1 – 2013/2515 =20 percent is probably a more accurate measure of the gender wage gap.

[16] 'Urban China' in Table A.4 and throughout this paper refers to China's largest cities – specifically the four municipalities directly under the jurisdiction of the central government (Beijing, Shanghai, Tianjin and Chongqing) plus the 15 subprovincial cities.

[17] Rapid educational upgrading since the 2005 Census also implies that Table A.4 is likely to overstate the education gap between the XMRC ads and Xiamen's 2010 labor force.

to the currently employed population. Finally, while definitional differences make it very hard to compare occupations on XMRC and the Census, it appears that jobs in production, construction and manufacturing are under-represented on XMRC, while professional and technical jobs are highly over-represented. Again, this is consistent with XMRC's focus on skilled workers, a population we know is less subject to gender profiling than less-skilled workers.

## 1.3 Gender Matching, Compliance and Enforcement: Aggregate Statistics

Aggregate statistics on applications and callbacks are shown in Table 1.1, broken down by the three job types in our data: jobs requesting women (F jobs), jobs requesting men (M jobs) and jobs that do not state a gender preference (N jobs). Turning first to total gender matching, row 1 shows the share of callbacks that are female ($\delta$) by job type. These statistics indicate a high congruence of the callback pool with employers' stated requests. Specifically, 94.0 percent of callbacks to F jobs are female and 100-3.7=96.3 percent of callbacks to M jobs are male. Combining F and M jobs, 94.8 percent of callbacks to gendered job ads are of the requested gender. Row 2 shows the share of applications to the three job types that are female ($\alpha$). It suggests that applicants' compliance with employers' gender requests plays a substantial role in accounting for this high level of gender matching, since applicant pools are almost as highly sorted by gender as callback pools. Specifically, 92.6 percent of applications to F jobs are female and 100-7.9=92.1 percent of applications to M jobs are male. Combining F and M jobs, 92.5 percent of applications to gendered job ads are of the requested gender.

The remaining rows of Table 1.1 show that employers' enforcement of their own stated requests also helps to account for the overall amount of gender matching that occurs. Specifically, in jobs explicitly requesting female applicants, men who apply are only 1/1.246=80.3 percent as likely to be called back as women. In jobs requesting men, female applicants are only 44.5 percent as likely to be called back as a man. Thus, at least in the raw data, employers' enforcement of their own gender requests is stronger against women applying to male jobs than men applying to women's jobs.

To get a better sense of the overall amount of gender matching and its components, it is useful to define the following index of gender matching:

$$G = \frac{g - g_0}{1 - g_0} \tag{1.1}$$

where $g$ is the share of gendered ads that are of the requested gender and $g_0$ is the share of gendered ads that would be of the requested gender if there was no gender matching (i.e. if we re-allocated the total population of called-back workers across all jobs – whether F, N and M – so that the total number of callbacks to each job remained the same, but the gender mix of callbacks was equalized across all jobs). Thus $G = 1$ if all callbacks to gendered jobs match the employers' request, and $G = 0$ if the female share of callbacks ($\delta$) equals its population average in all jobs. In our data, $g = .948$ and $g_0 = .501$, so our overall index, $G = .897$. In other words, on a scale where zero indicates no gender matching and 1 indicates perfect matching, the total amount of matching equals 0.9. With this index in hand, we can assess the relative contributions of compliance and enforcement to gender matching, $G$, using the identity:

$$\delta^J = \frac{\theta^J \alpha^J}{\theta^J \alpha^J + (1 - \alpha^J} \tag{1.2}$$

where $J = F, N, "or" M$ and $\theta$ is women's relative risk of being chosen from the applicant pool, i.e. the ratio of callback rates (f/m). Equation (2) allows us to compute two counterfactual levels of $g$ and $G$.[18] *Counterfactual 1* (no compliance) keeps enforcement, $\theta$, at its actual level in each of the three job types, but sets $\alpha$ (the share of women in the applicant pool) at its population mean level in all jobs (i.e. at .541, from Table 1.1). *Counterfactual 2* (no enforcement) keeps compliance, $\alpha$, at its actual level in each job type, but sets $\theta$ (women's relative risk of being picked from the applicant pool) at its population average, .866, in all jobs. The results are reported in Table 1.2.

According to row 2 of Table 1.2, eliminating worker compliance while maintaining actual levels of enforcement would reduce the share of callbacks that are of the requested gender, $g$, from .948 to .617. The corresponding decline in the gender matching index, $G$, is from .897 to .232. Thus, workers' compliance with employers' gender requests accounts for (0.897-0.232)/0.897=74 percent of the gender matching in our data. According to row 3, eliminating employers' enforcement while maintaining actual levels of worker compliance would have a much smaller impact, reducing g from .948 to .921 and $G$ from .897 to .842. Thus, active enforcement by employers of their own gender requests accounts for only (.897-.842)/.897=6 percent of the gender matching in our data. Because the decomposition in equation (2) is exact but nonlinear, the remaining 20 percent of gender matching is due to the interaction between compliance and enforcement.[19] We conclude that *compliance, i.e. applicants' self-sorting according to employers' gender requests in job ads, accounts for the vast majority of gender matching in gendered ads*. The intuition is straightforward: Because applicant pools are so highly

[18] Note that the G index depends on the relative sizes of the three job types (J), as well as on the overall share of workers who are called back to each job type. Throughout the paper, we design our counterfactual thought experiments to hold both of these quantities constant, varying only the gender mix of workers who apply to different job types (or firms, occupations, etc.) and the gender mix of callbacks.

[19] By 'exact' we mean that eliminating both compliance and enforcement would reduce G to zero.

gender-segregated, even completely equal treatment of male and female applicants in all job types would have only a small impact on the gender mix of callbacks to each job if application patterns are held fixed.

To put our estimates of gender-matching, compliance and enforcement in context, Appendix A.4 presents comparable measures of those three quantities for employers' gender, age, education and experience requests, as well as for the match between the posted wage and the applicant's current wage (when reported). Thus, for example, only 43.6 percent of call-backs, and 44.4 percent of applications match the employer's education request. (An education match means the worker's education falls into the category –primary or less, junior middle school, high school, college/technical school, or university—that is requested by the employer.) This compares to 94.8 and 92.5 percent for gender matching.[20] More broadly, compliance, enforcement and total matching are all greater for gender than for these other four characteristics (though age comes a close second under some measures). While these differences are particularly dramatic on the worker self-selection side, substantial enforcement differences are also present: The shares of age-, education-, experience- or wage-mismatched applicants that are called back all exceed 25.2 percent, compared to 5.2 percent of gender-mismatched applicants. Together, these statistics suggest an especially important role for gender, relative to these other characteristics, in determining what employers and employees consider to be a good match.

[20]Mismatch in education, experience and wages is measured by the indicators used in Table 1.4's callback regressions, which are based on broad categories. Additional details are provided in Table A.5. Table A.5 also presents results for a variety of measures of age-matching, as robustness checks.

## 1.4 Regression Analysis — Compliance

Section 2's aggregate statistics exhibit a high apparent level of worker compliance with employers' explicit gender requests: according to Table 1.1, $F, N$ and $M$ job ads attract applicant pools that are 92.6, 44.7 and 7.9 percent female respectively. Depending on which types of jobs explicitly request men and women, these large differences could over- or understate the causal effect of attaching an explicit gender label to a typical ad. For example, if gender requests are primarily used as a type of affirmative action (i.e. to attract workers to jobs in which their gender is underrepresented), these raw gaps would underestimate the causal effects of explicit labels on application behavior. Helleseter et al. (2020), however, show that explicit gender labels mostly reinforce prevailing stereotypes; thus Table 1.1's raw statistics could substantially overstate the causal effect of attaching a gender request to a job ad.

To adjust for these confounding factors, this Section takes two complementary approaches. In the first, we regress the female share of applicants to an ad on explicit gender requests, with controls for a detailed list of skill requirements and other desiderata, plus firm and job title fixed effects. Job titles are the main heading in every job ad. They provide a brief description of the job and can run up to 18 words on XMRC. For example, here is a random sample of ten (translated) job titles on the XMRC website: front desk administration assistant, project engineer, quality control, shift leader, customer service maintenance specialist, administration, ME product engineer, experienced two-dimension designer, customer service engineer, and front desk clerk. Job titles provide considerably more relevant information about the type of work than even the most granular standardized occupational classification systems. For example, Marinescu and Wolthoff (2020) found that job titles on Careerbuilder.com were much more predictive of advertised wages than 6-digit SOC codes, and were essential controls for identifying the effect of advertised wages on the number and quality of applications an ad received. Thus, in this approach we will be comparing the gender mix of applications to observationally identical ads for a very narrowly defined type of work, holding constant the identity of the firm advertising the job.

In our second approach, we replace the job title fixed effects in the above analysis by indicators of the predicted, or implicit 'maleness' or 'femaleness' of the job derived from a machine learning analysis of the words in the titles. Essentially, we use the words in the title to predict whether a person reading it can infer whether the job is likely to request men, or to request women. While these two predicted probabilities ($M_p$ and $F_p$, respectively) absorb less variation in job characteristics than the full set of title fixed effects, they provide a simple structure that helps us identify the types of jobs where inserting a gender label into a job ad has the largest estimated impact on application behavior. Notably, in both our estimation approaches in this Section, we use the entire sample of job ads available to us, not just the subset for which callback behavior is observed. To check for robustness, we replicated both analyses for the 'callbacks' subsample with very similar results.[21]

### 1.4.1 Approach 1: Job Title Fixed Effects

As noted, here we run regressions in our entire sample of 42,744 ads, where the dependent variable is the share of applications that are female ($\alpha$).[22] The regressors of

[21] Appendix Table A.6 reports these results for the title-fixed-effects approach.

[22] Appendix A.5 shows that requesting men (women) reduces the total number of applicants by 28 (31) percent. This is consistent with the idea that firms who post gender requests are choosing to restrict their attention to a smaller applicant pool (Kuhn and Shen, 2013). Gender requests appear to have no effects on the mean education and experience of the applicant pool, or on the share of applicants who satisfy the job's experience, education and age requirements.

interest are the labels attached to the ad ($F, N$ or $M$). In more detail, we estimate:

$$\alpha_j = a + b_1 F_j + b_2 M_j + cX_j + e_j \tag{1.3}$$

where j indexes jobs (ads), $F(M)$ is a dummy for whether the job requests women (men) and $N$ is the omitted job type. In column 1 of Table 1.3, we include no controls ($X_j$). Column 2 adds controls for the following job characteristics: requested education, experience, and age; the advertised wage; a dummy for whether a new graduate is requested; the number of positions advertised; plus dummies for missing education, age, wage and number of positions. Columns 3-5 in turn add occupation, job title and firm fixed effects, and column 6 interacts these job title and firm fixed effects. Thus, column 6 compares applicant pools across ads posted by the same firm for the same detailed job title, but with different gender requests. The extent to which the $b_1$ and $b_2$ coefficients attenuate as we add these controls captures the extent to which explicit gender labels are correlated with other features of job ads (such as a typically male occupation or job title) that allow applicants to infer the ad's desired gender even in the absence of an explicit gender request.

Table 1.3 shows that, as expected, the unadjusted effects of both the M and F job labels attenuate substantially – from 35 to 15 percentage points for M labels and from 50 to 25 percentage points for F jobs – as we add detailed controls for job and firm characteristics. Essentially all of this attenuation results from adding controls for occupation and job titles in columns 3 and 4 respectively: different types of work attract different ratios of men and women, most likely because men and women train for different types of duties and may have different preferences. In contrast, adding firm effects in column 5, and interacting them with job titles in column 6 has almost no effect, suggesting that detailed job duties are gendered in very similar ways by different employers.

This noted, the estimated effects of the gender labels remain economically large and highly statistically significant even in column 6, which compares the same job title in the same firm with different gender labels attached. Specifically, adding a request for men raises the male share of applications by 14.6 percentage points, or 26.4% (on a base of 1-.447 from Table 1.1); adding a request for women raises the female share of applications by 24.6 percentage points, or 55.0%.

It is worth noting that column 6's estimates are not driven by a single large firm, job title or title*firm cell: the 1,448 job ads that identify column 6 represent 416 distinct job titles posted by 505 different firms, and comprise 686 title*firm cells. In Appendix A.4 we show that these 1,448 ads are very similar to the full sample on most characteristics (including education, age, and experience) though they advertise somewhat lower wages (11%). Four of the five most common broad occupations in the identifying sample (construction, sales, administration and manufacturing) are in the top five overall, and nine of the most frequent *job titles* in the identifying sample are in the top ten overall. Notably, while some firms request both men and women for the same job title at different times, most of the 'gender-request-switching' that occurs within firm*job title cells takes the form of either switching between F and N requests, or between M and N requests. In other words, for a substantial number of job titles, firms sometimes request a particular gender, and fail to make a gender request at other times. This is the main source of variation that identifies the M and F coefficients in column 6 of Table 1.3. Finally, Appendix A.4 shows that estimates of column 6 that leave out one job title at a time are all very close to the full-sample estimates. Together, these patterns suggest that adding an explicit gender request to a job ad has substantial causal effects on the gender mix of applications it will receive. In other words, employers' gender requests appear to direct workers' applications.

### 1.4.2 Approach 2: Implicit Maleness and Femaleness

To better understand the source of the apparent compliance effect identified in Table 1.3, we now try to identify the types of jobs in which making an explicit gender request has the largest effects on application mix. If prospective applicants are using gender labels and other features of the job ad to predict whether a person of their gender would have a good chance of receiving a callback, we would expect explicit requests to have the largest impact on applications in *jobs where it is difficult for workers to infer the employer's gender preferences from the other contents of the ad*. To formalize this notion, we now replace the job title fixed effects in Table 1.3 by predicted probabilities that the job requests men (women), calculated from the words that appear in the title. Treating each ad's job title as a document, we calculate the implicit maleness and femaleness of each job using the Bernoulli naïve Bayes classifier (McCallum et al., 1998) for document classification; classifiers of this type are widely used in predicting whether a document is of a given type, for example a spam email. This methodological innovation solves a common problem in the analysis of job board data, namely how to collapse the high degree of granularity with which jobs are described into lower-dimensional measures that are theoretically relevant to a research question. In our case it allows us to summarize the implications of job titles for the employer's likely gender preferences with two scalars, summarizing the job's expected 'maleness' and 'femaleness'.

Briefly – details are available in Appendix A.6 – for each word, w, that appears in our entire set of job titles, we first estimate the probability of observing that word in the title of a job that requests men, $Prob(\text{"}observe\ word\text{"}\ w \mid \text{"}job\ requests\ men\text{"})$ using empirical frequencies. Next, treating job titles as 'baskets of words' which appear independently, we can compute the probabilities of observing a given job title, $k$, given the job requests men, $Prob(\text{"}observe\ title\text{"}\ k\ |\text{"}job\ requests\ men\text{"}\ )$ from its constituent words. Finally, using Bayes formula plus an assumption about workers' prior beliefs, we can compute the predicted maleness of each job title based on the words it contains.[23] Using the same procedure to predict each title's femaleness yields the two continuous variables,

$$M_p \equiv Prob(\text{"}job\ explicitly\ requests\ men\text{"}\ |\ \text{"}jobtitle\text{"}\ k) \tag{1.4}$$

$$F_p \equiv Prob(\text{"}job\ explicitly\ requests\ women\text{"}\ |\ \text{"}jobtitle\text{"}\ k) \tag{1.5}$$

which we use in our empirical analysis to represent the information contained in the job title about whether the job is likely to request men or women. Overall, $M_p$ and $F_p$ are quite predictive of employers' actual requests, with correlations of .411 and .402 with actual requests for men and women (which are binary variables) respectively. As we might expect, $M_p$ and $F_p$ identify what we might think of as stereotypically male and female jobs: the five 'most female' job titles (starting with the highest) are "front office desk staff", "administration office staff", "office staff", "cashier" and "administration assistant". The five 'most male' are "driver", "technician", "warehouse managing staff", "warehouse manager", and "production manager".[24] These indices of implicit maleness or femaleness allow us to estimate the effect on application behavior of adding an explicit gender request to jobs that 'look the same' to workers in terms of an employer's likely gender preference, and to see in which types of jobs the effect of explicit requests on application behavior is the greatest.

[23] We adopt the naïve prior that the unconditional chances a job requests men equals 50 percent. This simplifies the computations and reflects the idea that individual jobseekers may not have access to good summary statistics on the share of jobs of different types available to them.

[24] Additional examples of job titles at different levels of $F_p$ and $M_p$ are provided in Figure A.2.

More specifically, we now regress the female share of applicants to a job, $\alpha_j$, on employers' explicit gender requests (F and M), plus all the control variables used in column 5 of Table 1.3 (other than the job title fixed effects) plus quartics in the implicit maleness or femaleness of the job that workers could infer from the job's title ($M_p$ and $F_p$). In addition, each of these quartics is interacted with the three explicit job types, F, N and M.[25] These interactions allow, for example, the effect of an explicit request for women to differ in jobs that are stereotypically male (based on the words that appear in the job title) from jobs whose titles do not convey an obvious gender preference.

Predicted male and female applicant shares from these regressions are shown in 1.1. Part (a) of the Figure shows the predicted female applicant share as a function of the predicted femaleness of the job based on the words in the job title, separately for the three types of jobs (F, N and M). Predicted maleness is held fixed at its mean. Part (b) is the corresponding figure for male applicant shares as a function of perceived maleness, holding predicted femaleness at its mean. Finally, part (c) shows the estimated effects of encountering a request for a particular gender (relative to a non-gendered job) on the share of that gender in the applicant pool, with 95 percent confidence bands. These are the distances between the top two curves in parts (a) and (b).

Figure 1.1 shows, first of all, that explicit requests for male and female applicants have stronger effects on the gender mix of applications when the words in the job title do not send clear signals about whether the employer is likely to prefer men or women (i.e. when $M_p$ and $F_p$ are low). For example, when $F_p$ is near zero, the predicted effect on the female applicant share of inserting an explicit request for women into an N job is about 53 percentage points. This effect diminishes to about 26 percentage points when $F_p$ equals 0.7. A similar pattern is present for men, though it is less pronounced.[26]

[25] Appendix A.4 replaces this quartics by other functional forms, specifically a logistic specification, and dummies for each quartile of $M_p$ and $F_p$, and finds very similar results.

[26] We also note, however, that explicit gender requests continue to have substantial and highly signif-

Second, there is a subtle but interesting gender difference regarding when explicit requests matter. In 'not-obviously-female' (low $F_p$) jobs, women comprise a relatively large share of applicants only when the job explicitly requests women. In 'not-obviously-male' (low $M_p$) jobs, men comprise a relatively large share of applicants both when men are explicitly requested, and when the ad does not make a gender request. Together these patterns help us understand the much larger impact of F labels than M labels on the applicant mix in Table 1.3. Essentially, the main gender difference in application behavior occurs in jobs that – based on their title – are neither stereotypically male nor female. If we think of applying for jobs as entering a competition to get hired, these patterns are evocative of well-known gender differences in entry into competition (Niederle and Vesterlund, 2007), and of gender gaps in the propensity to apply for jobs in the presence of ambiguity (Gee, 2019).

We conclude our discussion of compliance effects with a reminder that our substantial estimated compliance effects are consistent with at two very different underlying mechanisms. One is that job labels communicate information about a worker's chances of getting a callback; in this view, women avoid male jobs because they know they have a lower chance of getting those jobs if they apply. The second mechanism is that – much like labels on men's and women's clothing—job labels communicate information about whether the worker is likely to want the job, without conveying any reluctance by the firm to transact with the worker. In this mechanism, women avoid male jobs because women dislike certain job attributes – perhaps competitive pay policies, long and inflexible hours, or even the absence of female co-workers – associated with those jobs. Assessing the relative importance of these two mechanisms requires an analysis of how gender-mismatched applications are treated when they are made,

icant effects on application behavior at all levels of $F_p$ and $M_p$. It follows that the jobs' implicit maleness or femaleness would not be close substitutes for explicit gender requests, if those requests were prohibited.

which is our goal in the next Section.

## 1.5 Regression Analysis — Enforcement

Section 2's aggregate statistics suggest a substantial amount of apparent enforcement by employers of their own explicit gender requests: according to Table 1.1, conditional on applying, women's callback rate in explicitly male jobs is 4.3 percent, compared to 8.7 percent in non-gendered jobs – a mismatch penalty of 4.4 percentage points, or 51 percent. Men's callback penalty from applying to explicitly female jobs, defined analogously, equals 9.0 - 5.8 = 3.2 percentage points, or 36 percent. Depending on which types of workers decide to apply to gender-mismatched jobs, however, these differences could over- or understate the change in callback chances that a representative worker would experience if she redirected her application from a non-gendered job to an identical job that requested the opposite gender.

To see this, imagine first that (say) women who apply to jobs requesting men are better qualified on dimensions like education, experience, and unobserved ability that the applicants hope will compensate for being of the 'wrong' gender. For the same reason, women may restrict their applications to jobs that fit their qualifications more closely when applying to explicitly male jobs. In both these cases, workers who make gender-mismatched applications will be positively selected on unobservables, and Table 1.1's raw mismatch penalties will underestimate the adverse effects of gender mismatch on the callback rate (because the people who choose to cross-apply are better-qualified and better matched than those who do not).

Alternatively, selection into mismatch can be negative, for example, if the women who apply to jobs requesting men are less able, or apply to jobs more indiscriminately. This could happen because those workers have low application costs, are highly motivated to find a job, or are simply careless. In this case, Table 1.1's 4.4 percentage point mismatch penalty for women will overestimate the adverse effects of gender mismatch on the callback rate. Adding controls for worker qualifications and job-worker match should attenuate the magnitude of the estimated penalty towards its true, smaller value.

To distinguish between these scenarios – and thereby measure just how 'hard' or 'soft' employers' explicit gender requests are – , we run linear probability regressions in a sample of applications, where the dependent variable is an indicator for whether the worker received a callback. In doing so, we control as tightly as possible for other aspects of match and worker quality that might affect callback rates. Of particular note, we control for unobserved worker ability by using worker fixed effects – i.e. we will compare the callback rates of the same worker who sends her resume to two observationally-identical jobs that differ only in their explicit gender label. We control for the detailed type of work using job title fixed effects. To account for the fact that people who apply to gender-mismatched jobs might be better or worse matched to the job on dimensions other than gender, we also include detailed controls for matching on a variety of characteristics.

In more detail, we estimate the following linear probability model:

$$Callback_i = \ alpha + \beta_1 FtoF_i + \beta_2 FtoM_i + \beta_3 MtoF_i + \beta_4 MtoM_i + \delta MWorker_i + \phi X_i + \epsilon_i \tag{1.6}$$

where $i$ indexes applications. Of the six possible application types, women applying to nongendered jobs ("F to N" ) is the omitted type. In this specification, $\beta_1$ and $\beta_2$ give the effect on women of applying to "M" and "F" jobs (relative to nongendered jobs), while $\beta_3$ and $\beta_4$ give the effect on men of applying to "M" and "F" jobs (again, relative to nongendered jobs). The parameter  gives the callback gap between men and women applying to nongendered jobs. Our main focus will be on the *gender mismatch penalties* associated with applying to a job that is targeted at the 'other' gender, $\beta_2$ and $\beta_3$.

Column 1 of Table 1.4 estimates equation 6 without controls, replicating the unadjusted gaps in Table 1.1. Column 2 adds controls for the job's requested education, experience and age; the advertised wage; and an indicator for whether a new graduate is requested. Also included are indicators of the match between the applicant's characteristics and those requirements, including indicators for whether the applicant's education, age and experience are below or above the requested level, the match between the advertised wage and the applicant's current or previous wage, and the match between requested and actual new-graduate status. Column 3 adds controls for the following worker (CV) characteristics: whether he/she attended a technical school; the applicant's *zhicheng* rank; whether an English CV is available; the number of schools attended, experience spells and certifications reported.[27] Indicators for applicant height, myopia and marital status are also included, all interacted with the applicant's gender.[28]

Column 4 adds fixed effects for the occupation of the advertised job, using XMRC's occupational categories. Column 5 adds job title fixed effects plus two indicators of the amount of competition for the job: the number of positions advertised and the number of persons who applied to the ad.[29] Our most saturated specification is column 6, which adds a full set of worker fixed effects. In this case, the effects of fixed applicant characteristics ("detailed cv controls" and the main gender effect) are no longer iden-

[27] Zhicheng is a nationally-recognized worker certification system that assigns an official rank (from one through six) to workers in almost every occupation. Ranks are based on education, experience and in some cases nationwide or province-wide exams.

[28] These 'detailed CV controls' introduced in column 3 are not requested in job ads very often, so it is not practical to construct variables summarizing their match with the job's requirements.

[29] These 'queue length' or 'submarket tightness' controls account for the possibility that overall competition for callbacks might be systematically stiffer in some job types than others. For example, callback rates in jobs that request women might be lower for all applicants if women 'crowd into' those jobs more than men crowd into jobs that request men (Sorensen, 1990).

tified, but our main coefficients of interest – which are interactions between job and applicant gender – can still be estimated. In effect, column 6 compares the outcomes of the *same worker* who has applied to observationally identical jobs that differ only according to the gender label (F, N or M) attached to the job, while allowing for this effect to differ according to the applicant's gender.

Before discussing our main coefficients of interest, it is worth noting that whenever they are statistically significant, observable indicators of the match between worker qualifications and job requirements are of the expected signs in Table 1.4: workers who have less education or experience than requested, or are older than requested are less likely to be called back. Finally, the job competition controls (not shown) are always highly statistically significant, indicating that these highly localized measures of labor market tightness have strong effects on the chances of being called back. Also of some interest, workers with more education than the job requests also experience a statistically significant callback penalty in all specifications but one(Shen and Kuhn, 2013).

Turning to the mismatch penalties, both men's and women's penalties attenuate somewhat as we add covariates in Table 4. As discussed, this pattern suggests that gender mismatched applicants are negatively selected, perhaps because they are less discriminating in where they send their applications. Despite this attenuation, however, the estimated mismatch penalty remains economically and statistically significant in the presence of worker fixed effects (column 6). For a woman, applying to a job requesting men reduces her callback chances by 3.7 percentage points, only a little less than the unadjusted effect (4.4 percentage points). For men, the attenuation is more pronounced – from 3.3 to 2.2 percentage points – suggesting more negative self-selection. In Appendix A.4, we probe this negative selection hypothesis further by examining the application behavior of workers who make gender-mismatched applications. We find that gender-mismatched applications come from workers who submit more than twice as many applications, compared to gender-matched applications. Gender-mismatched applications also go to a much wider variety of occupations and job titles, and tend to go to occupations and job titles with lower mean callback rates. Finally, gender-mismatched applications are significantly less likely to satisfy the job's age and education requirements.[30]

Summing up, our preferred estimates in Table 1.4 (column 6) imply that both men and women face substantial callback penalties when they apply to jobs that request the 'other' gender. While our estimates do not support the hypothesis that being of the requested gender is an essential requirement to get a callback, they do imply that applicants who choose to apply to gender-mismatched jobs pay a price in terms of a lower chance of getting a callback. Notably, this price (at 3.7 percentage points, or 43 percent) is higher for women than men (2.2 percentage points, or 24 percent), a difference which is highly statistically significant.

Two potential concerns with the above estimates are the possibility of gender misclassification and the effects of luck in the application process. Concerning gender misclassification, if some workers' genders are miscoded in their XMRC profiles our estimates of mismatch penalties would likely be underestimates, since some apparently gender-mismatched applications might be revealed as gender-matched on closer inspection by the employer. To check for this, we searched our data for individual workers who apply to an unusually large number of apparently gender-mismatched jobs, and excluded them from our sample. Appendix A.7 shows that excluding workers who direct more than half of their applications to opposite-gender jobs has almost

[30]The one exception to this pattern is that cross-gender applications are more likely to meet or exceed the job's experience requirements. While this might reflect conscious positive selection on this one dimension, it could also result from the fact that the average resume has much more experience than the average job demands (3.23 versus 1.13 years, respectively).

no effect on the results.[31]

Concerning luck, our results could overstate employers' openness to gender-mismatched applicants if a significant number of mismatched applicants are called back only because no candidates of the preferred gender applied to the job (Lang et al., 2005; Lazear et al., 2018). While our job competition controls capture some of these effects, a more direct test is to look directly at applicant pools containing zero applicants of the requested gender. As it happens, none of the 666 male jobs in our dataset received zero male applicants. We did find five female jobs that received no female applicants, and these jobs did call back some men. However, these jobs constitute less than 0.6 percent of the 867 female jobs in our sample.

We conclude this Section with two important caveats regarding the interpretation of our enforcement estimates. The first is that the our estimated mismatch penalties in callback rates do not in themselves constitute evidence for any particular form of discrimination, such as taste-based or statistical discrimination. Indeed, mismatch penalties are consistent with a number of underlying processes, including gender differences in productivity (both real and imagined) and the tastes of employers, recruiters, co-workers and customers, with the important proviso that any such productivity or taste differences must be highly *job-specific* to explain the patterns in our data: men need to be strongly preferred in some jobs, and women in others. To distinguish among these possible sources of mismatch penalties, research needs to examine the precise types of jobs in which they occur. For example, to assess the role of job-specific productivity differences one could look at tasks where there is established evidence of gender differentials in performance (Baker and Cornelson, 2018; Cook et al., 2021). Customer tastes

[31]Miscoding of the requested gender is not a concern since our data are the exact record of requested gender that workers observe on the job board when deciding where to apply. See Appendix A.7 for additional discussion of how gender is coded on the job board and on how we construct our "gender misclassification-robust" subsample of applications.

could be isolated by looking at jobs involving customer contact, and at employers' requests for applicant beauty. Indeed, Helleseter et al. (2020) find some support for a customer-tastes explanation of a significant share of explicit gender requests. Specifically, they find a large group of ads requesting young, attractive women in customer-contact jobs.

A second caveat concerns treatment effect heterogeneity. Specifically, while we have a number of controls for the quality of the match between the worker and the job, it is important to remember that our estimates still represent *treatment-on-the-treated* effects on the sample of applications people choose to make to gender-mismatched jobs. If workers disproportionately apply to the gender-mismatched jobs where they know their personal gender mismatch penalty (i.e. their personal treatment effect) is small, our estimates in Table 1.4 will underestimate the callback penalty associated with a randomly-selected gender-mismatched application.

## 1.6 Discussion

We believe that this is the first paper to study how workers respond to a common practice in developing-economy labor markets –explicit gender requests in job ads—and also the first to study how employers treat applicants to these types of ads. Our best estimates suggest that gendered job ads direct workers' applications away from jobs requesting the 'other' gender, and that employers penalize workers who apply to gender-mismatched jobs (in the form of a lower callback probability). Our estimated mismatch penalty is substantially greater for women who apply to men's jobs than for men who apply to women's jobs.

To assess some additional implications of our estimates for men's and women's financial wellbeing, in Appendix A.4 we analyze the wages advertised by jobs requesting men and women. Controlling for both firm fixed effects and job title fixed effects (i.e. holding fixed both the job's detailed duties and the individual employer's tendency to pay above or below the market) we find that all three job types (F,N, and M) request essentially identical amounts of education and experience. Jobs that request women, however, offer wages that are 192 yuan/month or 7.8 percent less than both N and M jobs, a difference which is highly statistically significant.[32] Thus, by directing women away from M and into F jobs, gendered job ads are also directing women into lower-paying jobs.

Since gendered job ads mostly direct men and women into gender-stereotypical jobs, they may also have consequences for gender segregation across occupations, firms, and individual jobs (ads). Assessing the effect of gendered ads on segregation from our estimates, however, is not possible without strong assumptions. With that caveat in mind, Appendix A.8 uses the following assumptions to estimate the effects of banning gendered job ads:[33]

(a) The causal effect of adding a gender request to an otherwise-unchanged job ad on the gender mix of applicants is given by Column 6 of Table 1.3.

(b) Conditional on applying, the relative callback rates of female applicants in all jobs ($\theta$) is unaffected by an ad ban.

(c) Segregation within each of the three job types (F,N and M) is unaffected by a ban: removing, say, all the female requests on the board does not affect workers' choices among the jobs that were formerly labeled as female.

[32] Recall that the unadjusted wage gap between M and F jobs, without these controls, is considerably higher, at 20 percent (from Table A.1).

[33] The United States banned gendered job ads in 1973, following a U.S. Supreme Court decision (Powell Jr, 1972). Austria effectively banned them in 2004, as part of the Austrian Equal Treatment Act. See Walsh et al. (1975)for a fascinating study of gendered job ads in the United States prior to the 1973 prohibition.

(d) Employers cannot circumvent the ban, for example by using code words and other signals of their gender preferences to direct applications.

(e) The ban does not change the types of human capital workers choose to invest in (for example, men's decisions to train as nurses, or women's as electricians).

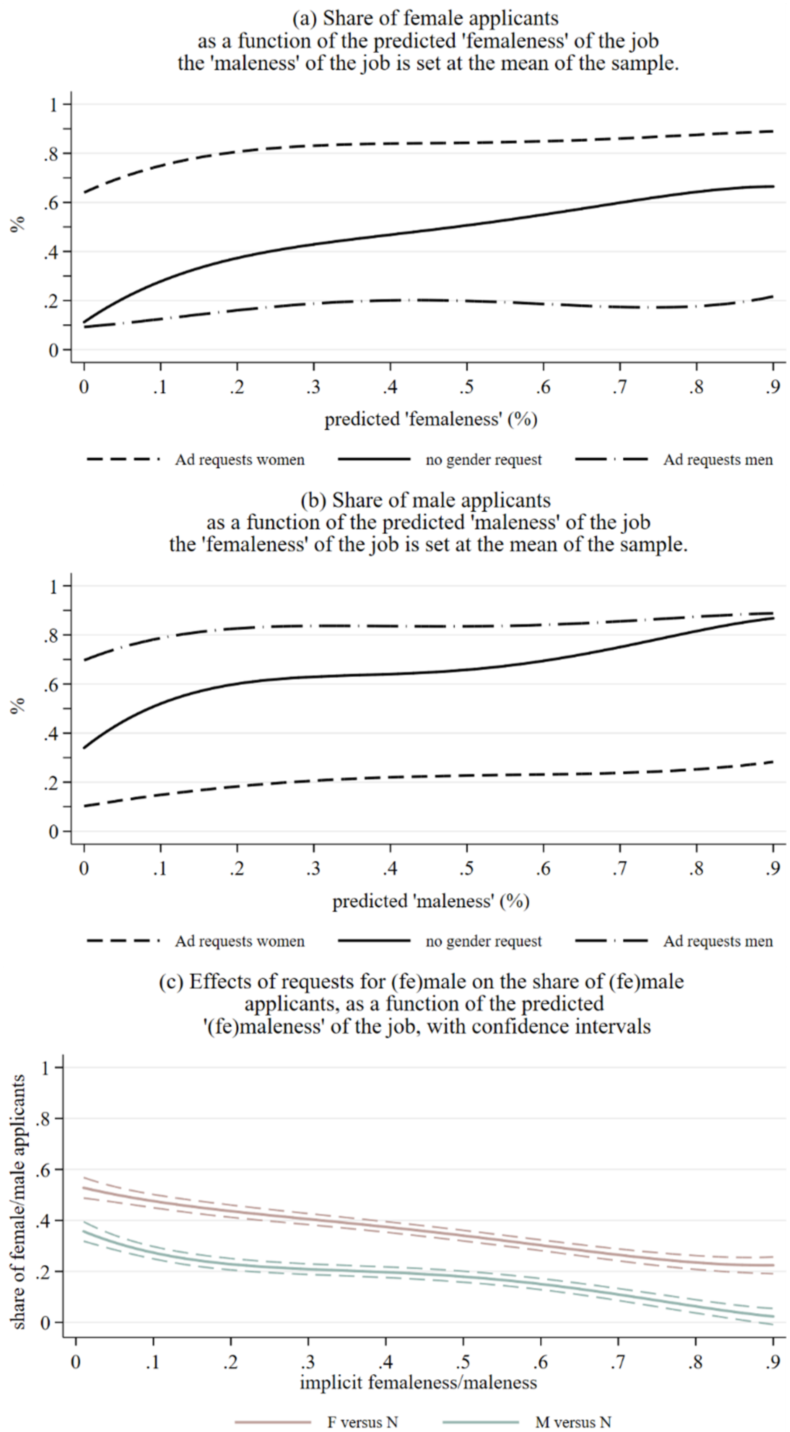
Under these assumptions, we estimate that banning explicit gender requests on XMRC would reduce gender segregation across jobs, firms and occupations by about 28, 27 and 19 percent respectively. While these findings are quite robust to changes in assumption (b), other changes are much harder to assess. Most importantly, we caution that the actual reduction in segregation could be much smaller if (d) was violated, and much larger in the long run if (e) was violated.

Because all our results are from a single job board, and because our estimates of causal connections are not based on random assignment, we view our analysis as the first rather than the last word on the effects of gendered job ads in labor markets. In our view, further analysis could profit from work in at least three different directions. First, it would be of interest to conduct a resume audit study of employers' 'enforcement' decisions: how will employers respond when we send identical resumes of different genders to jobs that request male versus female applicants? We view such an analysis as complementary with our internal job-board-based approach, because resume audits typically achieve tighter identification at the expense of focusing on only a handful of jobs. This is a significant issue in an environment such as ours, where employers' gender preferences vary dramatically across jobs.

Second, it would be useful to conduct a natural field experiment (Leibbrandt and List, 2015; Ibañez and Riener, 2018) on workers' compliance decisions: How, if at all, do workers' application decisions change when they are exposed to identical job ads that differ only in the presence or absence of a gender request? Again, such an approach would be complementary with a job-board-based approach because it provides a better-identified estimate, but for a small subset of jobs.

Finally, internal job-board data could be fruitfully used to study natural experiments associated with the imposition of a gendered-ad ban. An appealing feature of this approach is that it would allow investigators to study the simultaneous changes in both worker behavior (compliance) and firm (enforcement) behavior that result from such a ban. In addition, to the extent that a job board constitutes a local, occupational or national labor market, such a study would capture general equilibrium effects of the policy change, none of which are addressed by the preceding approaches.

Figure 1.1: Effects of Gender Requests and Implicit Gender of the Job Ads on the Gender Composition of Applications Received



Notes:

1. Figures represent predicted values of the female/male share of applicants ($\alpha$) from a specification identical to column 5 in Table 1.3, where the job title fixed effects are replaced by quartics in $F_p$ and $M_p$, each interacted with explicit job type (F, N and M).

2. Predictions in part (a), which shows the effect of implicit femaleness ($F_p$), hold $M_p$ at its mean. Predictions in part (b), which depicts the implicit maleness ($M_p$), hold $F_p$ at its mean. All other characteristics are set at their means. The regression is weighted by the number of applications to each ad, and standard errors are clustered at the firm level.

3. Part (c) shows the predicted effects of attaching an explicit male (female) label to a job ad (relative to an N label) at different levels of implicit maleness (femaleness), with 95 percent confidence bands. Notably, both effects are larger in jobs whose title does not convey a clear preference for the applicant's gender. In addition, the effects of explicit requests for women on application behavior are significantly larger (both economically and statistically) than the effects of explicit requests for men.

4. Predictions for values of $F_p$ or $M_p$ greater than 0.9 are imprecise and not shown; only 2,462 ads have values in this range, comprising .0377 and .0330 of the sample respectively.

Table 1.1: Application and Callback Patterns by Job Type

| | **Ad Requests Women** <br> ***F* jobs** | **Gender not specified** <br> ***N* jobs** | **Ad Requests Men** <br> ***M* jobs** | **All Ads** |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Share of callbacks that are female ($\delta$) | 0.940 | 0.437 | 0.037 | 0.505 |
| Share of applications that are female ($\alpha$) | 0.926 | 0.447 | 0.079 | 0.541 |
| women's callback rate ($f$) | 0.072 | 0.087 | 0.043 | 0.078 |
| men's callback rate ($m$) | 0.058 | 0.090 | 0.096 | 0.090 |
| ratio of callback rates ($\theta = f/m$) | 1.246 | 0.958 | 0.445 | 0.866 |
| $N$ of ads | 867 | 2,104 | 666 | 3,637 |
| $N$ of callbacks | 4,859 | 11,569 | 2,817 | 19,245 |
| $N$ of applications | 68,638 | 130,266 | 30,712 | 229,616 |

Table 1.2: Actual and Counterfactual Gender-Matching Rates

| | **Share of callbacks that are of the requested gender ($g$)** | **Gender-matching index ($G$)** |
|---|---|---|
| | (1) | (2) |
| **Baseline:** Actual values | 0.948 | 0.897 |
| **Counterfactual 1, no compliance:** Equal female share in applications ($\alpha$) in all jobs | 0.617 | 0.232 |
| **Counterfactual 2, no enforcement:** Equal female callback advantage ($\theta$) in all jobs | 0.921 | 0.842 |

Notes:

1. The population female applicant share ($\alpha$ ) (.541) is applied to all three job types when calculating counterfactual 1.
2. The population female risk ratio ($\theta$) (.866) is applied to all three job types when calculating counterfactual 2.
3. The gender matching index is calculated as $G = \frac{g-g_0}{1-g_0}$, where $g_0 = .501$.

Table 1.3: Effects of Gender Requests on the Share of Female Applications Received ($\alpha$)

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Ad requests men ($M$) | -0.3547*** | -0.3226*** | -0.2459*** | -0.1222*** | -0.1203*** | -0.1462*** |
| | (0.006) | (0.006) | (0.005) | (0.005) | (0.005) | (0.021) |
| Ad requests women ($F$) | 0.4954*** | 0.4519*** | 0.3736*** | 0.2263*** | 0.2339*** | 0.2462*** |
| | (0.005) | (0.005) | (0.005) | (0.004) | (0.005) | (0.023) |
| Primary School | | 0.0247** | 0.0095 | -0.0019 | -0.0057 | -0.0292 |
| | | (0.011) | (0.009) | (0.005) | (0.006) | (0.022) |
| Middle School | | -0.0627*** | -0.0507*** | 0.0036 | -0.0055 | -0.0343 |
| | | (0.011) | (0.011) | (0.006) | (0.007) | (0.027) |
| Tech School | | 0.0673*** | 0.0477*** | 0.0004 | -0.0014 | -0.0415** |
| | | (0.008) | (0.007) | (0.005) | (0.005) | (0.020) |
| Post-secondary | | 0.1159*** | 0.0639*** | -0.0016 | -0.0061 | -0.0408* |
| | | (0.008) | (0.007) | (0.004) | (0.005) | (0.023) |
| University | | 0.1203*** | 0.0499*** | -0.0137** | -0.0125* | -0.0189 |
| | | (0.010) | (0.008) | (0.006) | (0.007) | (0.037) |
| Number of positions advertised | | -1.7400*** | -0.9615*** | -0.1220 | -0.1338 | -0.5756 |
| | | (0.164) | (0.121) | (0.124) | (0.130) | (0.479) |
| Occupation Fixed Effects | | | Y | Y | Y | Y |
| Job Title Fixed Effects | | | | Y | Y | |
| Firm Fixed Effects | | | | | Y | |
| Title*Firm Fixed Effects | | | | | | Y |
| $N$ (ads) | 42,744 | 42,744 | 42,744 | 42,744 | 42,744 | 42,744 |
| "Effective" $N$ | 42,744 | 42,744 | 42,744 | 25,438 | 23,819 | 1,448 |
| $R^2$ | 0.554 | 0.590 | 0.721 | 0.925 | 0.950 | 0.974 |

Standard errors in parentheses, clustered by firm. *** p<0.01, ** p<0.05, * p<0.1

Notes:

1. In addition to the covariates shown, columns 2-5 also control for the following job ad characteristics: requested experience level (quadratic), requested age level (quadratic in midpoint of range), advertised wage (quadratic in midpoint of bin; 8 bins), dummy for whether new graduate requested, number of positions advertised, plus dummies for missing education, age, wage and number of positions.

2. All regressions are weighted by the total number of applications received.

3. 'Effective' N excludes job titles, firm IDs, and title*firm cells that only appear in one ad in columns 4, 5 and 6 respectively.

Table 1.4: Effects of Gender Requests on Callback Rates

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Female Worker * Female Job | -0.0149 | -0.0105*** | -0.0101*** | -0.0098*** | -0.0136*** | -0.0153*** |
| | (0.009) | (0.002) | (0.002) | (0.002) | (0.002) | (0.003) |
| Female Worker * Male Job | -0.0440*** | -0.0425*** | -0.0423*** | -0.0410*** | -0.0326*** | -0.0371*** |
| | (0.013) | (0.004) | (0.004) | (0.004) | (0.006) | (0.008) |
| Male Worker * Female Job | -0.0328*** | -0.0271*** | -0.0272*** | -0.0208*** | -0.0215*** | -0.0216*** |
| | (0.010) | (0.003) | (0.003) | (0.003) | (0.004) | (0.005) |
| Male Worker * Male Job | 0.0054 | 0.0016 | 0.0017 | 0.0038* | -0.0055 | -0.0155*** |
| | (0.009) | (0.002) | (0.002) | (0.002) | (0.004) | (0.005) |
| Male Worker | 0.0038 | 0.0004 | -0.0029 | -0.0065*** | -0.0173*** | |
| | (0.006) | (0.002) | (0.002) | (0.002) | (0.002) | |
| Education less than requested | | -0.0055** | -0.0047* | -0.0070*** | -0.0081*** | -0.0095*** |
| | | (0.002) | (0.003) | (0.003) | (0.002) | (0.004) |
| Education more than requested | | -0.0048*** | -0.0084*** | -0.0069*** | -0.0014 | 0.0020 |
| | | (0.001) | (0.002) | (0.002) | (0.002) | (0.003) |
| Age less than requested | | -0.0005 | -0.0018 | -0.0020 | -0.0036* | -0.0020 |
| | | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| Age more than requested | | -0.0330*** | -0.0309*** | -0.0284*** | -0.0205*** | -0.0215*** |
| | | (0.003) | (0.003) | (0.003) | (0.003) | (0.004) |
| Experience less than requested | | -0.0062*** | -0.0066*** | -0.0080*** | -0.0094*** | -0.0070*** |
| | | (0.002) | (0.002) | (0.002) | (0.002) | (0.003) |
| Experience more than requested | | 0.0004 | 0.0020 | 0.0012 | -0.0013 | 0.0013 |
| | | (0.002) | (0.002) | (0.002) | (0.002) | (0.004) |
| Wage below advertised | | -0.0010 | -0.0008 | -0.0020 | -0.0001 | -0.0015 |
| | | (0.002) | (0.002) | (0.002) | (0.002) | (0.003) |
| Wage above advertised | | 0.0009 | 0.0007 | 0.0002 | -0.0060*** | -0.0045 |
| | | (0.002) | (0.002) | (0.002) | (0.002) | (0.003) |
| Detailed CV controls | | | Y | Y | Y | |
| Occupation Fixed Effects | | | | Y | Y | Y |
| Competition Controls | | | | | Y | Y |
| Job Title Fixed Effects | | | | | Y | Y |
| Worker Fixed Effects | | | | | | Y |
| 'Effective' $N$ | 229,616 | 229,616 | 229,616 | 229,616 | 229,590 | 192,681 |
| $R^2$ | 0.001 | 0.005 | 0.005 | 0.016 | 0.198 | 0.388 |

Notes:

1. In addition to the covariates shown, columns 2-6 include the following controls for ad characteristics: requested education (5 categories), experience (quadratic), age (quadratic), the advertised wage (quadratic in midpoint of bin; 8 bins) and a dummy for whether a new graduate is requested. Columns 2-6 also include a dummy for whether the applicant's new graduate status matches the requested status, plus indicators for missing age and wage information for either the ad or the worker.

2. "Detailed CV controls" (used in columns 3-6) are an indicator for attending technical

school; the applicant's zhicheng rank (6 categories); an English CV indicator; the number of schools attended, job experience spells and certifications reported; and the following characteristics interacted with gender: height, myopia, and marital status (interacted with applicant gender)

3. Occupation fixed effects control for the 37 categories used on the XMRC website.

4. 'Effective' N excludes job titles and worker IDs that only appear in one ad in columns 5 and 6 respectively.

# Chapter 2

# Should I Show or Should I Hide – When Do Jobseekers Reveal Their Wages?

## 2.1 Introduction

A large literature has studied firms' incentives to voluntarily disclose verifiable private information that is relevant to their customers, workers and investors. This literature includes theoretical models (Jovanovic, 1982; Grossman, 1981; Milgrom, 1981) and empirical studies of firms' decisions to reveal information like financial performance (Depoers, 2000), HMO service quality (Jin, 2005) and restaurant hygiene (Bederson et al., 2018). This literature has also investigated the effects of public policies that mandate information disclosure by firms, such as impending layoffs (Kuhn, 1992), restaurant inspection reports (Jin and Leslie, 2003), the salaries of the other workers in a firm (Baker et al., 2019; Kim, 2015; Bennedsen et al., 2019), and the nutritional content of packaged food and restaurant meals (Mathios, 2000; Bedard and Kuhn, 2015).

During the past decade, a number of policies that regulate the disclosure of information about *workers* in labor markets have captured the attention of economists and policymakers. Interestingly, in contrast to mandating disclosure of the seller's characteristics, these policies aim instead to *restrict* the flow of verifiable private information from sellers (i.e. workers) to buyers. These policies include 'blinding' employers to the worker's gender or ethnicity (Goldin and Rouse, 2000; Krause et al., 2012; Behaghel et al., 2015), and prohibiting employers from asking workers about their criminal history (Agan and Starr, 2017; Doleac and Hansen, 2020), about their credit history (Bartik and Nelson, 2016; Bos et al., 2018; Ballance et al., 2020) and about their salary history (Agan et al., 2020; Hansen and McNichols, 2020; Khanna, 2020).

While workers, just like firms, have the option to voluntarily disclose almost any information they wish –and while such disclosure is generally not prohibited by the above labor laws–, there appears to be very little existing economic research on workers' *voluntary* disclosure decisions in the labor market: What do workers tell firms about themselves, and what do they hide? Under what conditions is disclosure or concealment more likely? In part, such studies have been rare because data on the information that is exchanged between individual workers and firms during the recruitment process has been hard to obtain and codify. Recently, however, Agan et al. (2020) have surveyed workers about how they would react to salary information requests from firms, and Kreisman et al. (2021) have used a large national sample of resumes to study which aspects of their educational background workers report or omit.[1] Closely related, Agan et al. (2021) use resume audit methods to study how U.S. recruiters *react* to (fictitious) workers' wage disclosure decisions; we compare our results on this question to theirs in Section 6 and find they are surprisingly similar.

[1] Kreisman et al. (2021) can identify true, missing information because they can link online resumes to records of educational institutions. In related work, Cullen and Perez-Truglia (2020) study workers' decisions to share their current salary information *with each other*.

Resume audits, however, cannot investigate workers' voluntary wage disclosure decisions which are our main focus here.

In this paper we exploit data on over a million resumes posted on a large Chinese job board, Liepin.com, to study job seekers' decisions on whether or not to reveal their current or most recent salary to potential employers. Because workers must input their salary information to create a profile on the board, we see the actual wages of both the revealers *and* the concealers, allowing us to identify exactly which wages are revealed, and to measure the extent to which employers can infer a concealing worker's current wage from other available information, including her application behavior and the other contents of her resume. Using internal information on how workers' applications were processed by recruiters also allows us to describe the association between disclosed wages and workers' success in the job search process.

Knowing which workers voluntarily reveal their wages to potential employers is useful for at least three reasons. First, identifying the revealers tells us which workers are most likely to be affected by labor market policies that limit recruiters' use of a worker's current salary information, such as the salary history bans (SHBs) that have been adopted in the United States and Canada since 2016. In the case of SHBs, Agan et al. (2020) argue that workers who volunteer their wage information without being asked ('always disclosers' in their model) are the most likely to experience spillover *benefits* from a salary history ban, possibly at the expense of the ban's intended beneficiaries.[2] If this group is disproportionately male and high wage, a salary history ban *that does not prohibit voluntary wage disclosure* could raise the gender wage gap.[3]

[2]This is because the SHB removes from the group of disclosers the ones with the lowest possible benefits from disclosure—i.e. the ones with less of the desirable characteristic that is signaled by disclosure.

[3]While most SHBs are ambiguous or silent about workers' right to disclose voluntarily, some of them explicitly protect this right. For example, in California, *"If an applicant voluntarily and without prompting discloses salary history information to a prospective employer, nothing in this section shall prohibit that employer from considering or relying on that voluntarily disclosed salary history information in determining the salary for*

Alternatively, in other contexts (such as the job board we study) one could imagine a platform reconfiguration that discouraged or prevented workers from revealing their current salaries to employers.[4] Knowing which workers voluntarily disclose today tells us which workers would be directly constrained by such a policy.

Second, workers' disclosure decisions convey key information about the nature of information asymmetries in labor markets: While labor market models with private worker ability (Gibbons and Murphy, 1992; Oyer and Schaefer, 2010) and models of market-based tournaments (Waldman, 2013; DeVaro and Kauhanen, 2016) emphasize the fact that workers' current wages can signal their unobserved ability, monopsony-based models (Manning, 2003; Card et al., 2018), efficiency wage models (Lazear et al., 2016) and some on-the-job search models (Burdett, 1978) emphasize the direct, reservation-wage consequence of having a higher current wage: Hiring this worker will almost surely cost the employer more than a lower-wage worker.[5] Because high current wages are 'good news' to prospective employers in the former case (at least if workers and firms share in the surplus from the employment relationship) and 'bad news' in the latter case, knowing how employers react to high versus low wages sheds light on the relative importance of these two canonical forms of asymmetric information. It can also reveal the circumstances under which the *ability-signaling* or the *reservation wage* effects of a higher current wage is more important to real recruiters.

Finally, we note that most workers –unlike sellers in product markets who may sell the same product millions of times– search for new jobs only a few times in their lives.[6]

*that applicant."* (California Legislature, 2017).

[4] Barach and Horton (2021) study a policy like this on Upwork.com. Such a move on Liepin would move it closer to a typical U.S. job board, since salaries are rarely listed in online resumes in the United States.

[5] In the efficiency wage case, the worker will also be less motivated at any given (new) wage.

[6] Topel and Ward (1992) report a career average of 11 jobs per worker in the United States. While this number is likely much higher in certain high-turnover industries and skill groups, it does not seem unlikely for the highly skilled professionals who look for work on Liepin.

Furthermore, these forays into the labor market are occurring during a period of rapid institutional change, as a number of competing intermediaries, including job boards, job and resume search engines, and matchmaking services continue to update their differentiated products.[7] In this situation, it is not at all clear that workers' decisions on how best to 'sell themselves' online are made in the most effective manner. Indeed, most of the statistics that would be needed for workers to make informed decisions about issues like resume design and wage disclosure are simply not available to them. Thus, an additional contribution of the current paper is to give workers information about how other workers make their disclosure decisions and how employers react to those decisions. This may shed some light on the efficacy of workers' current practices, and could suggest some changes in workers' strategies and in platform design that could benefit workers.

Our first main finding is that –as in empirical product market studies like Depoers (2000), Jin (2005), and Bederson et al. (2018), workers' disclosure of their current wages is *incomplete*, with only 40% of applicants revealing their current wages to employers.[8] As is widely recognized in the literature, partial disclosure is inconsistent with the unraveling predictions of early theoretical models, such as Milgrom (1981), but is consistent with a variety of extensions that allow for disclosure costs (Jovanovic, 1982; Fishman and Hagerty, 1989), unsophisticated buyers (Hirshleifer et al., 2004;

[7] Ongoing changes include integrating job boards with resume processing software, and using machine learning to propose matches to both firms and workers.

[8] While data on the frequency of voluntary wage disclosure by workers appears to be scarce, a number of recent surveys have asked workers about whether, and when their current employer asked them to reveal their previous wage. For example, Hall and Krueger (2012) found that 47 percent of U.S. workers reported that their employers had learned their pay in their earlier jobs before extending the offer. Barach and Horton (2021) reports 29.4% of respondents to their survey were asked by their employers about their compensation history, and 82.6% among them reported the inquiry occurred before making a job offer. A survey conducted by Agan et al. (2020) suggests that 25 percent of respondents reported that the employer asked for their current salary at some point during the application process in the last job they applied to. PayScale (2017) reports 43 percent of respondents were asked about their salary history during the interview process.

Fishman and Hagerty, 2003), buyers with heterogeneous tastes (Hotz and Xiao, 2013), imperfect competition among sellers (Board, 2009), sequential competition (Guo and Zhao, 2009) and multiple dimensions of quality (Levin et al., 2009).

Second, we show that the absolute level of a worker's current wage has a statistically significant but economically trivial effect on the likelihood a worker will disclose her wage. Third –and in contrast– three different indicators of whether a worker's current wage is higher than a reasonably sophisticated employer might forecast for a non-disclosing applicant strongly predict workers' decisions to reveal their current wages in their resumes. This suggests that workers, on average, expect that revealing a high wage (relative to a low one) is good news to employers: High wages signal higher ability without unduly discouraging employers due to higher expected wage costs. Notably, heterogeneity across workers in a 'naive' form of overconfidence is not a good explanation of these patterns: naively overconfident workers should both apply aggressively (to jobs that pay much more than their current wage) *and* reveal their current wages. We find, instead, that workers who apply aggressively are less likely to reveal their wages than other workers, suggesting a strategic motive.

Fourth, for reasons that appear to be unrelated to the above strategic considerations, men are much more likely to disclose their current wages than women. Specifically, in all regression specifications –and regardless of whether we control for the applicant's actual wage– women are about 20% less likely to reveal their current wages than men. Furthermore, this gender gap is even larger among workers whose wages are unexpectedly *low* (relative to their resumes and application decisions). This pattern suggests that greater male overconfidence contributes to the gender gap in wage disclosure, though as noted overconfidence cannot account for the association between wages and disclosure.

Fifth, even though workers seem to try to communicate unexpectedly high wages by revealing them (and to hide unexpectedly low wages by concealing them), evidence on how firms process applications suggests that these actions by workers are largely ineffective: The (positive) effect of having an unexpectedly high current wage on all three of our indicators of a candidate's success in the recruiting process is essentially identical, *regardless of whether the candidate reveals her wage or not*.[9] This suggests that employers can infer these idiosyncratic components of individual workers' ability from other aspects of workers' resumes that are challenging for investigators to encode. Interestingly, Banfi and Villena-Roldan (2019) find a parallel result for the wages employers attach to jobs (for example when setting up a profile for the job ad): these wages appear to direct workers' search even when they are not displayed in the job ad. Thus, workers appear to be able to make similar inferences from the text of job postings as employers do from resumes.

Finally, regardless of whether workers' wages are unexpectedly high or low, disclosing them appears to modestly *reduce* the probability that an application succeeds in the candidate selection process. Detailed analysis of the sources of these effects suggests two distinct explanations: While low-wage applicants tend to avoid disclosing their wages, workers in the bottom two wage deciles who *do* disclose their wages experience sharply lower success rates, and receive lower wage offers. We therefore interpret these disclosures as mistakes that convey negative information about ability that employers cannot infer from other resume contents. Disclosing a high wage (relative to other applicants) *also* reduces success rates, reducing the success-rate advantage of having a high current wage substantially. Since we also show that disclosing high wages raises the mean offered wage conditional on success, we argue that these disclosure decisions may perform a useful function for workers, by signaling a commitment

[9]Our three measures of success, in order of increasing exclusivity, are whether a candidate is marked as 'suitable', whether his resume was downloaded, and whether he was marked a a recruiting 'target'. Only 4.8 percent of resumes were marked as targets)

not to accept low wage offers.[10]

We conclude this Section by summarizing our paper's contributions to four distinct literatures: on salary history bans (SHBs); on the effects of workers' wage histories; on voluntary quality disclosure decisions; and on asymmetric information in labor markets. Our main contribution to the literature on salary history bans is to identify the workers who are most likely to continue to disclose their wages in the presence of laws (like those in the U.S.) that only prevent employers from *asking*. As one might expect, we find that these *always-disclosers* are disproportionately male. Perhaps less expected, conditioning on gender they are not disproportionately high wage earners. Instead, the disclosers are disproportionately workers who adopt a *conservative* application strategy (i.e., focusing on jobs where most other applicants are currently paid less than they are).

Relative to the (very small) literature on workers' voluntary wage disclosure, we expand the scope of investigation beyond workers' stated disclosure *intentions* (Agan et al., 2020), and beyond studies based on platforms for temporary or part-time jobs (Barach and Horton, 2021), to a large, naturally occurring labor market spanning a wide set of industries and occupations. Both workers and firms on our platform are searching for long-term, full-time job matches, and so face much higher stakes when making their revelation, application, and hiring decisions than on platforms like Upwork. In addition, relative to both Agan et al. and Barach-Horton, –who focus on wage disclosure during the post-interview stage– we focus on wage disclosure decisions in an earlier stage of the recruitment process: resume design and applications. This change in focus forces us to confront –both theoretically and empirically– the fact

[10]Specifically, having a higher-than-expected wage rises the probability of being marked as a target worker (our most exclusive indicator of worker success) by 0.0019 to 0.0034, which is an increase of about 4.0% to 7.1% relative to the mean targeting rate. *Revealing* this high wage offsets that positive effect by about one-third.

that a worker's wage disclosure decisions can determine whether they receive a job offer or not. Since this remains a possibility at all stages in the recruitment process, our theoretical and econometric approach to disclosure decisions can also be useful for the study of disclosure decisions in later phases of the recruitment process as well.

Third, together with Kreisman et al. (2021)–who study education revelation decisions– we extend the large literature on voluntary quality disclosure to decisions made by *workers*, as suppliers of labor. As we demonstrate in the paper, disclosure decisions in labor markets may differ from product markets because (a) workers (sellers) participate only infrequently in the market, (b) each worker has only one unit to sell, (c) each buyer (vacancy) is usually seeking to buy just a single unit, and (d) search frictions limit the pool of applicants from which the buyer can choose. Under these conditions –rather than always preferring high- over low-quality workers– buyers (firms) might rationally prefer workers who are 'good but too good', with interesting implications for workers' optimal disclosure decisions.

Finally, we contribute to the large literature on asymmetric information in labor markets by providing evidence on the nature of the information contained in workers' current wages: Overall, –with the possible exception of workers with very high wages– our results indicate that employers perceive high wages among their job applicants as, on net, an advantage, not a disadvantage. This provides empirical support for models of labor markets in which firms share rents with workers, and in which the most important unobserved feature of workers (from firms' point of view) is not the level of workers' outside options but their unobserved ability.

The remainder of the paper is organized as follows. Section 2 summarizes the structure and results of a model of optimal wage disclosure in a static search setting, which is consistent with two of our main empirical results: disclosure increases with workers' current wages (relative to employers' expectations based on application behavior and resume contents), and firms prefer workers with higher current wages (again relative to expectations), as long as those wages are not too high. Section 3 describes Liepin.com and the data we have constructed from its internal records, and Section 4 describes our estimation approaches. Section 5 presents our evidence on workers' voluntary wage disclosure decisions, and Section 6 provides evidence on how firms respond to workers' current wages and disclosure decisions. Section 7 concludes.

## 2.2 A Model

In this Section, we describe the structure main results of a model of workers' decisions on whether to disclose their wages in the early stages of job search, for example by including their wage in their resume or job board profile.[11] A more detailed exposition and proofs of these results are provided in Appendix B.1. In our model, we think of workers as sellers who can choose whether to disclose their quality (wage), and employers as buyers who maximize profits by hiring workers based on their current wages and disclosure decisions. The model considers a 'one-shot' job market where a fixed number of job seekers send a fixed number of applications to a fixed number of jobs (or 'firms'), then each firm can offer a job to as many applicants as it wants. All the workers are ranked in terms of current wages, which have an increasing, one-to-one relationship with the worker's productivity in every firm. All the firms are ranked in terms of productivity as well, and it is efficient for abler workers to match with more productive firms.

Our model deviates from models of quality disclosure in product markets in at least three key respects: (1) each seller (worker) has only one unit of the good to sell; (2)

[11] Although framed as including wages in one's resume, the main insights of our model apply to wage revelation at any stage of the hiring process, so long as revealing a wage that is too high can jeopardize receiving a wage offer.

workers apply to and receive offers from a finite number of firms; and (3) offers are costly for firms to make. Under these conditions –rather than always preferring high- over low-quality workers– firms might rationally prefer workers who are 'good but too good', and workers will design their disclosure decisions knowing that their applications will be seen by multiple, heterogeneous employers. In general, these employers will hold different beliefs about a non-disclosing worker's wage (because employers do not know which other jobs their applicants applied to). A worker's decision on whether to disclose will then depend on the mix of jobs to which he expects to apply.

The purpose of the model is to demonstrate two main results. First, even though we assume that job seekers' productivity increases monotonically with their current wages, in almost all firms the expected return to offering a job to a worker is a non-monotonic (hump-shaped) function of the worker's current wage.[12] In other words, each vacancy has a range of applicant wages to which it makes job offers, and the vacancy does not make any offers to workers with wages outside that interval. The intuition for this result is based on the fact that workers expect to receive more than one offer, but can only accept one of them. Because of this (as in the academic labor market), some high-wage workers will be 'too good' for firms to extend offers to.

Second, consider the relationship between a worker's current wage (relative to the expectations of the firms she will apply to) and the worker's privately optimal disclosure decision. Our model predicts that workers with the lowest wages will conceal their wages, while workers in a range of wages above that will reveal their wages. Workers with the highest wages may reveal or conceal their wages, depending on the distribution of firm productivities. Stated differently, workers who intend to apply very *aggressively* (to jobs where they are among the lowest-wage applicants) should hide their wages, and workers who intend to apply less aggressively should reveal

[12] The only exception is the most productive firm in the market.

them. The *least* aggressive applicants –i.e. the ones with the *highest* wages relative to their co-applicants– might prefer to reveal or conceal their wages.

The intuition behind this second result is based on the two distinct effects of having a higher current wage: higher current wages result in higher wage offers, but also affect the chances of getting an offer. At low wage levels, increases in the worker's current wage raise both offered wages and offer chances– thus, they make disclosure unambiguously more attractive. At higher wage levels, however, additional wage increases can reduce offer chances; if this effect is strong enough, concealment could be optimal for the highest-wage (i.e. least aggressive) applicants. Interestingly, this pattern mirrors the counter-signaling result in Bederson et al. (2018), but is driven by very different factors.

Our empirical analysis in the remainder of the paper supports both these predictions in the following senses: a) in their resume processing decisions, firms appear to prefer workers with high current wages, as long as those wages are not too high; and (b) workers' wage disclosure rates increase with their current wages (relative to employer expectations), except at the very top of the relative wage distribution. That said, our empirical analysis also reveals some additional patterns that are inconsistent with this model. Specifically, while our model describes two things quite well –the relationship between workers' relative wages and their disclosure decisions, and the way firms respond to workers with different current wages– the data indicate that employers respond to workers' actual, current wages in almost exactly the same way, *regardless of whether workers reveal them or not*. This raises two interesting questions which we explore further in our empirical analysis: First, do workers misunderstand that firms can infer their wages even when they don't reveal them? And second, do workers' wage revelation decisions signal something other than the value of the worker's current wage?

## 2.3 Setting and Data

Our data are from Liepin.com. Founded in 2011, Liepin is the fourth largest online job board in China; its focus is on the high-end professional labor market.[13] Liepin has a typical structure for online job platforms: Employers post job advertisements and workers post resumes, and each party can search the other side of the market using a number of filters. In addition, firms' hiring agents can process applications and contact applicants in Liepin's online environment. The master dataset used to construct both our main regression samples sample was constructed as follows: First, we identified all the firms advertising on Liepin that had more than 100 employees, then we collected all the new job ads posted by these firms between March 1 and April 30, 2018. We then followed this inflow sample of job ads until June 30, 2018 to collect information on all the applications they received, and on the success of those applications.

In the rest of this Section, we describe the recruitment process on Liepin as it affects job *applicants*, job *ads*, and job *applications*. We describe the construction and composition of our two main regression samples: the sample of job *applicants* used to study workers' wage disclosure decisions in Section 5, and the sample of job *applications* used to study the relationship between disclosure decisions and workers' job search outcomes in Section 6.

### 2.3.1 Applicants

To view job postings and apply for jobs on Liepin, a job seeker must first set up an account and create a profile. To create a profile, the applicant is prompted to enter information on her birth date, gender, current or most recent annual wage, marital

[13]Liepin explicitly targets jobs and workers with before-tax annual salaries above 100,000 RMB in first-tier cities (Beijing, Shanghai, Guangzhou and Shenzhen), or with before-tax salaries above 60,000 RMB in other cities(100,000 RMB is about 1.6 times the national average wage in China in 2018).

status, working history, education level, employment status, current industry and occupation. In the case of wages, workers can check a box that makes this information invisible to employers.[14] Thus we (the investigators) know the applicant's wage, even when the employer does not. In the case of marital status, job seekers can choose from three options: single, married, and confidential. If the job seeker sets her marital status as confidential, neither we nor the employer know her actual marital status.[15]

The education variables in the job seeker's profile include the highest degree attained, whether the degree is Tongzhao, the rank of the university that granted the degree, and whether the university is included in the 985/211 Projects.[16] The job seeker is asked to select her current employment status from four categories: 1) unemployed, 2) employed but desires to switch to a new job (intensive search), 3) employed and looking for new working opportunities (moderate search), 4) employed and have no plan to switch jobs. The work history that we can observe is the job seeker's total years of work experience, the tenure and industry in the last two jobs. In addition to the information on previous jobs, the job seeker's career expectations are also recorded, which include her desired industry, occupation, city, and province, and her desired annual wage.

While job seekers can register on Liepin for free, they have the option to buy a gold membership, which promotes their resumes to a higher position in employers' search results; gold membership status is invisible to employers. Based on the resume's quality (i.e. education level, current annual wage, and work experience in large companies), Liepin classifies it as either *white-collar* or *elite*; elite status is only visible to em-

[14]See the wage box in Figure 2.1. Current wages for unemployed workers are the wages in their most recent jobs.

[15]See Appendix B.7 for more additional details about the disclosure of marital status.

[16]Tongzhao degrees (awarded to students who took the 'normal' Gaokao during high school) are sometimes viewed as more desirable than equivalent degrees earned by people who took the 'adult Gaokao' or 'self-learning Gaokao'. Project 985 and Project 211 universities are elite institutions sponsored by the central and local governments. There are 39 985 universities and 112 211 universities.

ployers (not workers). The platform also assigns a completeness score to a job seeker's profile, which is displayed in the resume and is visible to both the job seeker and potential employers.[17]

Table B.1 presents descriptive statistics on the 941,733 job seekers who submitted applications during our observation window. 63.4% of the applicants were men. The average job seeker was 31.8 years old with 8.79 years of work experience; 28.2% of job seekers reported themselves as not currently employed. Reflecting Liepin's focus on highly skilled workers, 80.3% of job seekers held a bachelor's degree or above, and 30.9% of them graduated from 985/211 Project universities. The average applicant had a current annual wage of 177,018 RMB (around 27,000 US dollars), which was about three times the national average wage in China in 2018.

Although we cannot separately verify the wages of the applicants in our data, we have four reasons to believe that most workers do not misrepresent the wages they enter into their profiles. First, the fact that Liepin allows workers to conceal their wages from prospective employers gives workers a way to protect this information without having to misrepresent their wages. Second, the fact that Liepin allows workers to report both a current and a *desired* wage also allows workers to signal a (high) target or reservation wage without misrepresenting their actual wage. Third, most employers in China will ask job seekers to provide certification of their most recent salaries (such as a bank statement) after workers accept the job offer; employers can withdraw offers to less-than-honest workers. Finally, Liepin's job recommendation algorithms are based on the information provided by job seekers. Even if job seekers conceal their current wages from employers, the wage information is still included in the website's job recommendation algorithms. Thus, misrepresenting one's wage could lead Liepin to

[17] In addition to the completeness of worker-supplied information, this score also reflects Liepin's authentication of the worker's name and contact information.

make inappropriate recommendations. These institutional factors motivate our decision to model wage disclosure throughout the paper as truthful revelation of *verifiable* information, rather than as costly signaling or cheap talk.

As an extension of the preceding data on applicant characteristics, we also compute –for each applicant– measures of the match between the characteristics of their current job and those of the job they say they are seeking. These indicators show that over 78.3% of job seekers would like to stay in the same city, and about half are seeking jobs in their current industry and occupation. On average, workers were looking for jobs that pay 19.5% more than their current job.[18] Website classification variables suggest that the average applicant registered her account about two years ago. 84.5% of applicants were classified as *elite* workers, and 11.6% of applicants had a gold membership.

The share of workers voluntarily revealing their current wages to recruiters is 40.0%, which is only slightly lower than the share of workers disclosing their *desired* wages, 44.6%. A slightly lower share of workers (36.6%) revealed their marital status. Relative to workers who withhold their wage information, wage disclosers are more male, younger, less experienced and less educated.

### 2.3.2 Ads

Job ads on Liepin consist of four sections, the first of which contains basic information including the job title, location, industry, occupation, and the number of subordinates to the position. While hiring agents must specify a wage range for each job, they can choose to hide this information from workers, in which case the posted wage will be listed as "negotiate with the employer". The second section of a job ad lists the job

[18]The categories used to describe city, province, main and sub industry, and main and sub occupation in Liepin's system are the same for workers' current jobs, workers' desired jobs, and the jobs advertised by employers. In total, there are 505 cities, 41 provinces (including overseas), 12 main industries, 52 sub industries, 55 main occupations and 753 sub occupations.

requirements, which can include gender, age, work experience and education; all of these except gender are visible to workers.[19] The third part is a detailed job description of at least 60 words. The fourth and final portion allows hiring agents to choose various settings for the post: for example, employers can decide how long to post this vacancy, and have the option to post an estimate of their average response time to applications.[20]

In March and April 2018, 19,264 firms posted 328,921 job ads that received at least one application. According to Table B.2, three fourths of job ads revealed the job's wage range to job seekers. Only a few jobs indicated a preferred age (3.8%) or a preferred gender (0.9%), but almost all the jobs had requirements for the worker's education (93.5%) and experience levels (87.2%). About 70 percent of jobs requested candidates holding a bachelor's degree or above, and the same share requested at least 3 years of work experience. On average, a job ad received 26 applications, and employers claimed that it took 4.3 days for hiring agents to give feedback to applicants. Descriptive statistics on the firms who posted these ads are presented in Table B.3. Firms in the median size category (with 100-1000 employees) accounted for 70.1% of all total job vacancies. Above half of the jobs were posted by private firms, and a typical firm had three hiring agents who were responsible for job posting and recruitment.

### 2.3.3 Applications

Once a job seeker has completed her profile and identified a desirable job, she can apply to it by clicking the "apply" button in the posting. This transmits her resume to

[19] Since 2016, Chinese labor law has prohibited employers and job boards from posting job ads containing explicit gender requests (see Kuhn and Shen (2021) for the recent history of these regulations). While rare, the gender requests on Liepin may serve as reminders to HR agents of the firm's preferences for the job. We do not know if they are used by Liepin's internal algorithms for employers' search results.

[20] The longest posting duration for a job is three months, and the job ad will be withdrawn by the website automatically afterwards. Figure B.5 shows the timeline of a job ad in Liepin.

the firm's hiring agent responsible for the job.[21] After a job application is made, the website recommends 10 similar jobs to the job seeker. Job seekers can then immediately apply to all of these jobs using Liepin's "batch apply" function, which is similar to the "select and purchase all" button in online shopping websites. About one quarter of the applications in our data came from this "batch apply" approach.[22]

During our sample period (March and April 2018), employers on Liepin received about 8 million applications from workers on the site. The first thing a hiring agent sees once applications arrive is a set of summary cards, which display very limited information about each applicant, including gender, education level, age, years of experience, location and company name of the current or recent job. No wage information about the worker is displayed on these cards. To see the wage (and the rest of the applicant's resume) the hiring agent must click on (i.e. *view*) the summary card; this happened in 41.3 percent (or about 3 million) applications.[23]

Since employers cannot see whether a worker has disclosed her wage until they view the resume, our analysis of employers' responses to workers' wage disclosure decisions in Section 6 of the paper restricts attention to the 3,542,049 applications that were viewed by hiring agents (i.e. the agent clicked on the summary card to see the resume). Descriptive statistics on this sample are provided in Table B.4.[24] Almost all the applicants satisfied the job's posted requirements for age, education and experience. In 74.7% of applications, the jobs' locations were consistent with the workers'

[21] Figure B.4 shows the timeline of an application submitted in Liepin.

[22] In deciding where to apply, job seekers can also access additional information by clicking the *job lens* button in the requirements section of the ad. This gives them (among other things) the platform's estimate of their hiring chances, based on the match between the worker and the job. Only 3.2 percent of job seekers clicked this button, however. Appendix B.3.1 provides additional information about the *job lens* service.

[23] Descriptive statistics on all applications and viewed applications are provided in Tables B.5 and B.4 respectively. These samples are quite similar, though as one might expect the viewed resumes were better matched to the job. Further, Table B.6 shows that both the wage level and the wage disclosure decision have no effect on the probability an application is viewed by recruiters.

[24] Descriptive Statistics for the full sample of 8,488,353 applications are presented in Table B.3.

desired location, while the fraction is lower for industries and occupations. Thus, consistent with Marinescu and Rathelot (2018*b*) it appears that workers' job search is more constrained by locations than by industry and occupation on Liepin.

Although we do not observe employers' recruitment or callback decisions, we can construct three indicators of the hiring agents' interest in each candidate from the actions the agents take on the site. These indicators are the outcome variables in our analyses of how employers react to workers' wages and disclosure decisions. Specifically, hiring agents can mark a resume as a *target* candidate or as *unsuitable* for the position. They also have the option of downloading and saving the resume. In increasing order, our three indicators of an application's success are therefore (a) the agent does *not* mark the resume as *unsuitable* (henceforth 'marked *suitable*'– 56.8 percent of applications); (b) the agent *saves* the resume (8.0 percent of applications), and (c) the hiring agent marks the applicant as a recruiting *target* (4.8 percent of applications).[25] More details on how hiring agents process applications can be found in Appendix B.3.1.

In contrast, our analysis of workers' decisions on whether to disclose their wages in Section 5 uses data for all the workers who submitted an application during this period. To calculate the expected wages of these workers from the point of view of the 'average' employer they applied to, we use data on all the applications each worker sent, irrespective of whether those applications were viewed or not. Descriptive statistics of this worker-level dataset are provided in Table B.1.

[25]There are some overlaps between these categories. For example, 41.1% of applications marked as targets were saved by hiring agents, and 24.7% of saved resumes were marked as targets.

## 2.4 Estimation Approach

### 2.4.1 Identification Issues

When employers care about a piece of information that is missing from a worker's resume, both theory and evidence suggest that the employer will try to infer that missing information (in our case, the worker's current wage) from other observables, including the remaining contents of the worker's resume (Agan and Starr, 2017; Doleac and Hansen, 2020). To decide whether to reveal or conceal their wages, workers who anticipate this behavior therefore need to ask themselves "what would the employer infer about my wage if I concealed it?".

To empirically investigate whether firms and workers on Liepin behave in this manner, we therefore need to do two things: First, we need to model how employers draw inferences about the wages of the non-disclosers in their applicant pool. Second, because job seekers on Liepin make their wage disclosures *ex ante* (i.e. in their *resumes*, before applying to jobs) for each worker we need to aggregate these expectations over all the jobs she eventually applies to. In a little more detail, for each of three alternative models of employer wage expectations, we will construct an indicator, $OverWage_{ij}$, for whether learning worker $i$'s wage would be a positive surprise to firm $j$, where "positive" means that $i$'s actual wage is higher than the firm predicted. For each worker $i$, we then aggregate firms' expectations over all the jobs the worker applied to (her *application set*) to derive a worker-level indicator, $HighWage_i$. $HighWage_i$ tells us whether, on average, worker $i$'s actual wage would be a positive surprise if it was revealed to all the jobs she applied to. Finally, we study how workers' revelation decisions relate to their $HighWage_i$ status, and how employers' recruiting decisions respond to the $OverWage_{ij}$ status and wage revelation decisions of the workers in their applicant pools.

While in principle we could propose any model of employer expectations we like, we note three key features of our choices here. First, they are simple and intuitive heuristics that could plausibly characterize the beliefs of HR agents who have limited time, cognitive resources, and information about the wage distribution of their applicants. Second, an expectations function that approximates how employers interpret resumes with hidden information should predict workers' disclosure decisions in a particular way. Specifically, workers' disclosure decisions should be much more sensitive to the *gap* between a worker's actual and predicted wage than to her actual wage alone. Intuitively, workers whose wages are equal to what an average employer in their application set would infer (from the fact that the worker applied and the contents of their resume) can neither gain nor lose from revealing, because revealing conveys no additional information. On the other hand, workers whose actual wage would be a surprise to the employer could either gain or lose a lot from revealing. As we shall see, all three heuristics we construct have this empirical property.

Third, conditional on any expectations function, our data allow us to test whether firms' and workers' behaviors are consistent with each other. Specifically, if workers believe that high wages are good news to firms –in the sense that firms prefer workers with high current wages to otherwise identical workers with lower current wages– and if workers are correct in those beliefs, we should see that workers with unexpectedly high wages should be more likely to *succeed* in the application process, *and* should be more likely to *reveal* their wages, compared to other workers. The opposite should occur if workers believed that high wages were bad news to firms (because they simply communicated a high reservation wage).

Before describing our three empirical models of employer expectation formation, we make three comments about challenges to the identification and interpretation of our estimates. Focusing first on our analysis of workers' wage disclosure decisions in Section 5, we emphasize that we do not view the coefficients we estimate there as the causal effects of randomly assigning $HighWage$ status to job seekers. Instead, as noted, we view both the choice of which jobs to apply to *and* whether to include one's current wage in the resume as jointly determined aspects of a worker's *application strategy* when she creates a profile on Liepin. While workers with different unobservables (including 'ambition') might well choose different application strategies, the goal of our disclosure regressions is to test whether these two components of workers' application strategies are consistent with each other in a cross section of workers. Specifically, workers who choose to apply *aggressively* will (by definition) be classified as $LowWage$ workers (because they are applying where most of the applicants are paid better than they are). If –as our data indicate– high current wages are good news to employers, we should expect these 'aggressive' job seekers to be more likely to *conceal* their current wages. This somewhat counterintuitive prediction distinguishes our findings from, for example, a model where overconfident workers not only apply aggressively but naively advertise their current wages to the market at the same time.

Turning now to our estimates of how firms respond to individual workers' wage disclosure decisions in Section 6, the identification challenge that seems of greatest concern is unobserved worker heterogeneity: Even though we have access to all the information about workers that HR agents see when they make their recruiting decisions, we are still limited in our ability to encode all these items in a way that captures everything that matters to recruiters (such as addresses, names of high schools, and other subtle indicators of expertise and social class). Thus, even when comparing observationally identical resumes, it remains possible that some unobserved worker characteristic –such as their level of 'ambition'– accounts for both their wage revealing behavior and their success in the recruitment process. Fortunately, we can address this issue by taking advantage of an intriguing feature of our data– about 2.2% of job seekers change their revelation decision during their job search on Liepin. While this raises the issue of why workers changed their disclosure decision, it allows us to use worker fixed effects to control for un-encoded aspects of workers' resumes that might be correlated with the workers' revelation decisions.

Our final observation about identification applies to our estimates in Section 6 of how firms respond to workers' current wages, *conditional* on workers' disclosure decisions. For example, we may wish to estimate how firms respond to a higher wage when the wage is visible. Here, we can take advantage of a unique feature of Liepin's platform– the 'batch apply' option, which automatically sends applications to ten board-selected jobs. Since workers who choose this option cannot de-select any of the jobs, this gives us the option of using only *within-batch* wage variation to study employers' responses to different current wage levels. Arguably, this within-batch wage variation is exogenous since the candidate is forced to apply to all of the jobs in the batch.

### 2.4.2 Modeling Employers' Estimates of Non-Disclosers' Wages

We conclude this Section by describing our three empirical models of how employers estimate the current wages of non-disclosing workers. To different degrees, the three approaches take advantage of three pieces of information a rational employer should consider when trying to guess the 'true' wage of a non-discloser: 1) the worker chose not to disclose, 2) the worker chose to apply to this job, and 3) the other contents of the worker's resume.

**Approach 1: $W_E^j$ = lower bound of job $j$'s posted wage**

According to this heuristic, employers who have posted a job, $j$, infer that the non-disclosing workers in the pool of applicants to job $j$ have a wage equal to the lower bound of the posted wage range for that job. One motivation for this heuristic derives from the unraveling predictions of early models of non-disclosure like Milgrom (1981): If higher current wages are viewed positively by the employer, then all workers but the lowest paid will want to distinguish themselves from workers below them by revealing, and firms will infer that non-disclosers have low wages. More broadly, we might expect disclosers to be negatively selected, on average. A different rationale (which is more in the spirit of our model) is the notion that workers who already earn close to the maximum offered for the advertised job have little to gain by applying to it. Thus we expect applicants' current wages to cluster toward the bottom end of each job's advertised wage range.

**Approach 2:** $W_E^j$ **= median**$(w_1^j, w_2^j \ldots, w_{mj}^j)$

According to this heuristic, the expected wage of non-disclosing workers applying to job $j$ is the median current wage of the workers who applied to it.[26] While in some sense this is a naive inference, we note that (unlike Approach 1) it relies on the notion that employers have some experience with jobs of this type: They know enough about the wages of non-revealers to guess at a median overall wage.[27]

**Approach 3:** $W_{E_i}^j = f(X_i \mid job\ j)$

In the previous two heuristics, employers based their wage inferences purely on the fact that the worker applied to the job, and not on the contents of his resume. In this heuristic, both sources of information are used. The expected wage of a worker with characteristics $X_i$ who has applied to job $j$ is modeled as the predicted wage from

[26] In our main analyses we used the median wage among all applicants (both concealers and disclosures). Robustness checks using only the revealed wages were almost identical– see Table B.20 and Table B.25.

[27] Notably, Approaches 1 and 2 both satisfy the ordering rule for expected wages in our theoretical model, $W_E^1 < W_E^2 < \ldots < W_E^M$. This is not necessarily true of Approach 3, because it allows firms' expectations to also depend on the individual workers' observable characteristics.

a regression of workers' previous wages on a set of demographics and characteristics *in the sample of applicants to job $j$*.[28] Details of these wage prediction regressions are provided in Appendix B.5.1.

A notable feature of all our employer heuristics is that different employers will make different inferences about the true wages of the same non-disclosing worker. This makes sense because the fact that a worker has applied to a job is informative about her true, current wage, and because employers don't see which other jobs their applicants applied to. Approach 3, however, is the only one that assigns different expected wages to different *non-disclosing* workers who apply to the same job. It does so because these workers' observables differ.

## 2.5 Results– Worker's Wage Disclosure Decisions

Since workers' disclosure decisions are made before applying to jobs on Liepin, studying those decisions requires us to aggregate employers' inferences about non-disclosers' wages across all the jobs that a worker expects to apply to.[29] We therefore start this Section by describing this aggregation process for each of the three preceding employer wage expectation models. Notably, all these aggregation methods are based on the set of jobs a worker actually applied to during our data window (i.e. the worker's *application set*); we use these jobs to assign each worker to either a *HighWage* or a *LowWage* category. Intuitively, these two categories summarize a worker's *application strategy*: *HighWage* workers are pursuing a *conservative* application strategy, because they are currently paid more than most of the other applicants at the jobs they apply

[28]We use mean rather than median predicted wages in Approach 3 to accommodate the low sample size in some of these job-level regressions.

[29]While workers can change the wage revelation decision in their Liepin profile, this is relatively rare: 2.2% of workers ever changed their wage disclosure decisions in the data window.

to. Thus, if these workers were to reveal their wage, most of the employers they apply to would be surprised at how high it was. By the same reasoning, *LowWage* workers are pursuing a *more aggressive* application strategy, which means that most employers they apply to would be surprised at how low their wage was if it was revealed.

### 2.5.1 Defining High-Wage and Low-Wage *Workers*

**Indicator 1:** $HighWage1_i = 1$ **if** $w_i >$ **median (lower bound of job's advertised wage range)**

If employers treat wage non-disclosers as if their current wage equals the lower bound of the job's posted wage range (Approach 1 to modeling employers' expectations), it seems reasonable to define a *worker* as high-wage ($HighWage1 = 1$) if his current wage exceeds the median of the posted lower wage bounds in his application set. Thus, high-wage workers are the ones who apply to jobs with disproportionately low lower-wage bounds. According to this definition, 51.3% of applicants are are high-wage workers.

**Indicator 2:** $HighWage2_i = 1$ **if** $w_i >$ **median**$(w_1^1, \ldots, w_{m1}^1, w_1^2, \ldots, w_{m2}^2, \ldots, w_1^K, \ldots, w_{mK}^K)$

If employers assume that non-disclosers earn the median wage of all applicants to the same job (Approach 2), it seems reasonable to define high-wage workers as having current wages that are above the median of the applicants who *ever* applied to the same jobs as the focal applicant did. In other words, for each worker, we first find all the applicants to the jobs that he applied to. Then we compare the worker's current wage with the median current wage of these applicants. The fraction of applicants who are high wage according to this definition ($HighWage2 = 1$) is 49.8%.

**Indicator 3:** $HighWage3_i = 1$ **if** $w_i > f(X_i|job\,1...K)$

If recruiters run 'mental wage regressions' to predict the wages of the non-disclosers in their applicant pools based on resume characteristics, it seems reasonable to use a similar approach to define *workers* as high- or low-wage in an *ex ante* sense. To that end, we classify applicants on Liepin as high wage ($HighWage3 = 1$) using the regressions described in Approach 3 above, but the regression sample is now all the workers who have applied to the same jobs as the focal applicant.[30] In other words, a worker is high-wage if his current wage is higher than one would predict from his resume, using data from workers who shared at least one job with him in their application sets. Using this method, 48.7% of applicants are high-wage workers.[31]

### 2.5.2 Determinants of Disclosure

We explore the correlation between workers' characteristics and their wage revealing decisions by estimating a linear probability model of the following form on all the 941,733 applicants in our sample:

$$y_i = \beta_0 + \beta_1 HighWage_i + AX_i + FE + e_i \tag{2.1}$$

where $y_i = 1$ if applicant $i$ reveals his current wage, $X_i$ is applicant $i$'s characteristics and $FE$ denotes various fixed effects, detailed below.[32] The coefficient of interest is $\beta_1$, which represents the effect of having an unexpectedly high current wage (relative to what employers would expect if the worker concealed it) on the job seeker's wage disclosure decision.

Table 2.1 shows estimates of equation (2.1) that successively introduce more de-

[30] Details about the wage prediction on worker level are shown in Appendix B.4.1

[31] We did not use the predicted wage from the regression on the whole applicant sample to reduce the prediction error.

[32] A small share of our workers –about 2.4 percent– changed their disclosure decision during our two-month sample period. For these workers, we used the wage revealing status that they used to send out the most applications as our indicator of their disclosure decision.

tailed controls for the applicant's characteristics. In column 1 the only covariate is an indicator for whether the worker is classified as $HighWage$. In column 2, we control for the worker's gender and the level of the worker's current wage. Column 3 adds controls for the worker's marital status and a quadratic in age. Column 4 adds detailed controls for the applicant's education and work experience: the highest degree obtained, whether the highest degree is Tongzhao, the domestic and world rank of the applicant's university, and whether that university is a 985/211 university; a quadratic in years of work experience; the applicant's employment status; and the applicant's industry and tenure in the last two jobs. Column 5 adds the following website classification variables: whether the resume is elite, how long the resume has been created, the score measuring the completeness of the worker's profile, whether the worker has a gold membership; and the total number of job applications sent during the data period. It also controls for the gap between the applicant's desired and current wage, and the match between the applicant's current location, industry and occupation and the desired ones. In column 6, we add fixed effects for the worker's current location, and column 7 adds fixed effects for the worker's current industry and occupation. We cluster standard errors in all specifications at the worker's sub-occupation level.

Regression results for our three high-wage indicators ($HighWage1$, $HighWage2$, and $HighWage3$) are presented in Panels A, B and C of Table 2.1. Across all three $HighWage$ measures and across all regression specifications, we find robust and statistically significant evidence that workers with higher-than-expected wages are more likely to disclose their current wages. In the most tightly controlled specification (column 7), the changes in disclosure probability range from 1.6 percentage points for $HighWage3$ to 5.1 percentage points for $HighWage1$, which represent increases of 3.9 and 12.6 percent relative to the average disclosure rate of 40.0%. Together, these results support an interpretation of the data where (a) high current wages are 'good news' to employers, and (b) workers whose actual wages are higher than their resumes suggest tend to disclose those wages for strategic reasons. This finding evokes similar results in product markets, where high-quality sellers are more likely to disclose (Mathios, 2000; Jin and Leslie, 2003; Jin, 2005), as well as survey results from Agan et al. (2020) who find that relatively highly paid workers are more likely to disclose.

To explore the relationship between the *size* of the potential surprise associated with learning a worker's wages and that worker's wage disclosure decision, Figure 2.2 plots the results of regressions that replace our binary $HighWage$ indicator with dummies for deciles of the $HighWage$ distribution. While the effects of a more positive wage surprise are robustly and monotonically increasing across almost all the deciles, two interesting additional features are evident. First, the largest disclosure effects are at the very bottom of the $HighWage$ distribution: Workers who have very low wages relative to employer expectations are much more likely to hide their wages than other workers. Thus, most of the effect captured by our binary $HighWage$ indicators comes from very low wage applicants (relative to expectations) choosing to hide that fact. Second, the probability of disclosing flattens out at about the 80th percentile of the $HighWage$ distribution, and begins to decline beyond that for two of our three $HighWage$ measures. As noted, this behavior is consistent with both the type of *countersignalling* found among the highest quality restaurants by Bederson et al. (2018), and with the predictions of our model for workers' disclosure decisions.

In contrast to the effects of having a wage that diverges from employers' expectations ($HighWage$), Table 2.1 shows that the *absolute* level of a worker's current wage has a trivial effect on the worker's wage revealing decision: earning 10,000 RMB (or 5.6 percent) more per year decreases the probability of revealing the current wage by .0004, which is a 0.1% reduction. We view this contrast as highly suggestive evidence in favor of our hypothesis that workers' wage disclosure decisions depend on the *new information* that might be revealed by doing so, not on publicly known factors like age, earnings, location, firm size, occupation and experience that affect wages in well known ways.[33]

Another intriguing feature of Table 2.1 is the highly significant, robust, and economically large effect of the applicant's gender on wage revelation decisions: Men are 7 to 9 percentage points (or about 20 percent) more likely to reveal their wages. Notably, the estimated size of this effect does not attenuate as we add detailed controls for resume characteristics, and all our estimates of the gender effect control for the job seeker's actual current wage, which is observed for all the job seekers in our sample.[34] This, together with the fact that a worker's absolute wage has a trivial effect on workers' disclosure decisions, rules out women's lower current wages as an explanation of their lower disclosure rates. Put another way, Table 2.1 suggests that gender has an independent effect on wage disclosure that is unrelated to the strategic factors we model here.[35]

Possible mechanisms for an independent gender effect on disclosure include the possibility that men and women have different psychological costs of competition, disclosing and bargaining (Stuhlmacher and Walters, 1999; Niederle and Vesterlund, 2007;

[33]Our results are also consistent with the research from Conlin et al. (2013). In their setting, students are free to choose to submit their SAT I scores or not when they apply to colleges. They find that applicants with higher SAT I scores are less likely to submit their score. Applicants who do worse than their fitted scores are likely to withhold their scores, and applicants with higher alternative measures of academic ability, like SAT II scores and high school GPAs are more likely to choose not to submit their score, all else equal.

[34]This gender disclosure gap is similar to the results of Conlin et al. (2013), who find that, when the submission of SAT scores is optional to college, women are more likely to not submit their SAT scores conditional on those scores, and with the survey result from Agan et al. (2020) who show that "Always-disclosers" (of current wages) are more male.

[35]Indeed, women's lower disclosure rates are quite hard to explain using these strategic factors: Suppose for example that women apply less aggressively than men do, and that for some reason our $HighWage$ controls do not completely capture this gender difference in application behavior. Then, on average, women will be disproportionately applying to jobs where they are better paid than the employer would expect, based on observables (including gender). If that was the case, women should disclose *more frequently* than men, in order to advertise the fact they are 'under-applying'.

Booth, 2009). Additional possibilities are that female applicants believe that prospective employers will underestimate their value, or that employers use the same low wage more aggressively as a bargaining tool when the applicant is female. While we cannot distinguish among these, one explanation we can partially rule out is a generalized aversion among women to disclosing personal information. This is because the marital status coefficients in Table 2.1 show that applicants who reveal their marital status are 1.5 to 1.8 percentage points *less* likely to reveal their current wages (see Appendix B.4.2).[36]

A final covariate of theoretical interest in the Table 2.1 regressions is the relationship between the worker's *desired* wage and the worker's disclosure decision. Consistent with our model, workers who are seeking the greatest wage gains in their next job (i.e. "aggressive" applicants) are less likely to reveal their current wages to prospective employers. Estimated effects of the remaining covariates in Table 2.1 are reported and discussed in Tables B.8 - B.10 in Appendix B.4.2.

### 2.5.3 Heterogeneity and Robustness Checks

In this subsection we explore how our main effect of interest –the effect of having an unexpectedly high wage on the probability a worker discloses her wage– varies across different types of workers. We begin by adding interactions between the high-wage measures and our *Male* indicator into the Table 2.1 regressions, to see if men and women respond to unexpectedly high wages differently. The resulting interaction coefficients –reported in Table B.11 in Appendix B.4.3– are significantly negative, indicating that (compared to men) women's disclosure decisions are *more* sensitive to having

[36] Zide et al. (2014), Jackson and Lilleker (2011) and El Ouirdi et al. (2015) review a large number of psychological studies of online disclosure differences between women and men, many of which find that women are less likely to disclose personal information than men.

an unexpectedly high wage. Quantitatively, among $LowWage$ workers (i.e. worker's whose actual wage would be a negative surprise) men are 11 percentage points more likely to reveal their wages than women. Among $HighWage$ workers, this gender gap shrinks to around 6 percentage points. In contrast to the overall disclosure pattern in our data (which suggests that, on average, $LowWage$ workers are rational enough to hide that fact from prospective employers) this pattern suggests a possible role for overconfidence, at least among among men relative to women (Niederle and Vesterlund, 2007): Men whose actual wages would be a negative surprise to employers are much less likely to hide that fact from employers than women are.

Since unemployed workers and on-the-job searchers have different job search patterns (Blau and Robins, 1990), we might expect their wage disclosure behaviors to differ as well. To address this question, we interacted our $HighWage$ indicators with an indicator for whether the applicant is currently unemployed and replicated our Table 2.1 regressions in Table B.12 in Appendix B.4.3. The significantly positive estimates of these interactions imply that high-wage unemployed workers have a higher propensity to disclose their (most recent) wages than high-wage employed workers. Unemployed workers who recently held a surprisingly well paid job might feel it particularly important to reveal that fact to prospective employers.

Although the regressions in Table 2.1 control for the worker's industry, occupation, previous wage, and education, it remains possible that the interesting and robust $HighWage$ effect in that table differs dramatically across these subgroups of workers, or is driven only by a small and particular set of occupations, industries or education levels. To explore these possibilities, we split the applicant sample into 12 industries, into the 15 most common occupations (accounting for 82% of workers), into 10 (absolute) wage deciles, and 7 education categories, and run Table 2.1's column 7 regression in each subgroup separately. Plots of the coefficients and 95% confidence intervals using the three high-wage indicators are displayed in Figures B.8 - B.11 in Appendix B.4.4. Reassuringly, the main results carry through to essentially all the subsamples: Compared to other workers, $HighWage$ workers are more likely to disclose their current wages across all industries, occupations, wage percentiles, and education levels.

Additional tests of the robustness of our Table 2.1 results –including additional definitions of $HighWage$ workers– are provided in Appendix B.4.4. Again, the main conclusion remains unaltered: Men disclose more than women do, and the more a worker earns relative to her expected wage, the more likely she will reveal her current wage.

## 2.6 Results–Employers' Responses to Wages and Disclosure

In the previous Section, we showed that workers behave *as if* higher wages were 'good news' to employer: workers reveal their wages when they are unexpectedly high and conceal them when they are unexpectedly low. In this Section, we shift our attention to employers to ask two questions: a) When employers can see workers' current wages, does a higher current wage increase employers' interest in hiring that worker?, and (b) How do employers respond to workers' decisions to reveal their wages, and how does that depend on whether those wages are unexpectedly high or low? To answer these questions, we return to the employer-side measures of wage expectations defined in Section 4 and classify each application an employer receives to job $j$ as either high- or low-wage. This gives us three alternative definitions of high-wage workers; to distinguish these job-level indicators from the previous Section's worker-level measures, we label them as $OverWage$ indicators. Specifically, applicant $i$ at job $j$ is classified as:

$OverWage1_{ij}$ if applicant $i$'s wage exceeds the lower bound of job $j$'s posted wage range.

$OverWage2_{ij}$ if applicant $i$'s wage exceeds the median wage of applicants to job $j$.[37]

$OverWage3_{ij}$ if applicant $i$'s wage exceeds her predicted wage from a regression among all the applicants to job $j$.

According to these three measures, 51.0, 49.8, and 47.8 percent of all applications were classified as $OverWage$ respectively.

### 2.6.1 Application Success Rates

To measure how workers' wages and their wage disclosure decisions jointly influence application outcomes, we estimate the following regression on the 3,542,049 applications in our data that were *viewed* by HR agents:[38]

$$\begin{aligned} Y_{ij} =&\alpha_0 + \alpha_1 OverWage_{ij} + \alpha_2(Disclose_{ij} \times UnderWage_{ij}) \\ &+\alpha_3(Disclose_{ij} \times OverWage_{ij}) + BX_{ij} + FE + e_{ij} \end{aligned} \tag{2.2}$$

The employer's responses, $Y_{ij}$, are binary variables derived from the way the HR agent processes a worker's application. As noted, we consider three different indicators of candidate success: $Target$, $Save$, and $Suitable$ (in diminishing order of exclusivity). In equation (2.2), $\alpha_1$ gives the effects of a higher-than-expected candidate wage on the success rates of candidates *who do not disclose their wages*. If none of the informa-

[37] As an alternative to this measure, we used only the applications that were *viewed* (i.e. clicked on) by hiring agents to construct this indicator. The results in Table 2.2 and Table 2.3 were very similar (see Table B.20 and Table B.25).

[38] Recall that a recruiter cannot see most of an applicant's characteristics (including the wage, or whether it is revealed) unless she clicks on the applicant's summary card on the recruiter's screen. Since we are interested in the effects of wages and wage revelation, we restrict our sample in this Section to *viewed* resumes only. A resume must be viewed before it can be saved/downloaded, marked as suitable, or marked as a recruiting target.

tion contained in our $OverWage$ indicators is available to employers unless the worker reveals them, $\alpha_1$ should equal zero. $\alpha_2$ measures the effects of *disclosing* a lower-than-expected wage on candidate success, and $\alpha_3$ measures the effects of *disclosing* a higher-than-expected wage. The effects of having a higher-than expected wage *among workers who disclose their wages* can be calculated from the above coefficients as $\alpha_1 + (\alpha_2 - \alpha_3)$; in other words if disclosure has the same effects on $OverWage$ versus $UnderWage$ workers ($\alpha_2 = \alpha_3$), then the effects of having a higher-than-expected wage are the same, regardless of whether the worker discloses it or not.

Table 2.2 shows the results from estimating equation (2.2) for $Target$, the most exclusive of our three indicators of candidate success (i.e. the closest to receiving a job offer). Results for the other two indicators are presented in Appendix B.5.2. Aside from some minor differences for the $Suitable$ indicator, which is by far the least exclusive of the three indicators, these results are very similar to the ones in Table 2.2.

Column 1 in Table 2.2 contains no controls, and column 2 controls for a variety of job characteristics. Column 3 adds controls for the match between the applicant and the job, such as whether the applicant satisfies the education and gender requirements, and column 4 adds controls for worker characteristics, such as current employment status, industry and tenure. Column 5 adds fixed effects for the job's location, industry, and occupation, plus fixed effects for the firm and the application date. Column 6 replaces the job characteristics with a job fixed effect. Finally, in column 7 we take advantage of the fact that a small share of workers (22,198 workers or 2.4 percent of our 941,733 applicants) switched their wage revelation decisions during our sample period.[39] While this raise the issue of why these workers decided to change their disclosure status, it provides an additional perspective by letting us compare the

[39]This total comprises 7,659 applicants who switched from concealing to revealing, 13,921 applicants who made the reverse switch, and 618 who applicants changed their wage disclosure choice more than once.

job application outcomes of the same worker when she makes two different revelation decisions. Standard errors are clustered by job.[40]

For all three of our indicators of $OverWage$ status and in all regression specifications, the $Overwage$ coefficient in Table 2.2 is positive, highly statistically significant, and economically substantial in magnitude. For example, in column 6 of panel A (the most saturated specification without worker fixed effects), having a higher-than-expected wage raises the candidate's chances of being marked as a recruiting target by 0.59 percentage points, relative to a mean targeting rate of 4.9 percentage points. This suggests that *even when candidates do not disclose their wages*, employers are able to infer the higher levels of productivity that are associated with higher wages from aspects of their resumes that we have not been able to encode in our control variables.[41] Employers' ability to infer productivity in this way is underscored by a second robust feature of Table 2.2: The $Disclose * UnderWage$ and $Disclose * OverWage$ coefficients are almost identical to each other across all specifications. As noted earlier, this means that the effects of having a higher-than-expected wage on firms' recruiting decisions are the same when firms can see worker's wages than when they cannot (because $\alpha_2 = \alpha_3$). Seeing a worker's wage thus appears to convey *no* additional information about productivity to employers.

To probe this idea further, Appendix B.5.6 uses the 'batch apply' feature of Liepin's website to test whether Table 2.2's main results might be driven by workers' endoge-

[40] Arguably, disclosure decisions made by switchers may be more endogenous than those of other workers, since the switchers are making conscious decisions to make a change.

[41] Consistent with this interpretation, we note that the $OverWage$ coefficient drops sharply in magnitude when we add applicant fixed effects. Interestingly, however, despite the lower magnitude, the $Overwage$ coefficient remains positive and statistically significant even in the presence of worker fixed effects. These remaining effects could occur because (a) Workers change other aspects of their resumes (in ways that are not captured by our covariates), and/or (b) workers change their application strategies (in ways that are not captured by our discrete $OverWage$ and $UnderWage$ indicators) when they switch to disclosing. For example, if workers switching to disclosure also switched to a slightly less aggressive application strategy (thereby raising their current wage relative to their competitors), this could account for the $OverWage$ effects in column 7 of Table 2.2.

nous decisions on where to send their resumes. When a worker uses this feature, he automatically applies to all ten jobs suggested to him by Liepin's algorithm. In Appendix B.5.6, we re-estimate Table 2.2 by restricting our estimation sample to batch applications and including a full set of batch fixed effects, we can eliminate all worker discretion in which jobs to apply to. Thus, we only use *within-batch* variation in whether the application was over- versus under-wage. Arguably, this within-batch variation in over-wage status is exogenous since the candidate is forced to apply to all of the jobs in the batch. We find that being a high-wage applicant increases the success rate in batch applications, but the penalty of disclosing a lower-than-expected wage is a little greater than disclosing a higher-than-expected wage.

A final main finding from Table 2.2 is that disclosing one's wage appears to *reduce* workers' application success rates, *regardless of whether the disclosed wage is high or low relative to the worker's co-applicants*. Again, the magnitudes are economically significant (about 0.2 to 0.3 percentage points relative to a mean of 4.8 percentage points). This striking finding raises two obvious questions: (a) what negative information is conveyed by disclosure *per se* that causes employers to shy away? and (b) why do *any* workers disclose their wages? To shed some light on these questions, Figure 2.3 explores how applicant success rates respond to a more continuous measure of $HighWage$ status. Specifically, Figure 2.3 displays the effects of ten deciles of $OverWage$ on $Target$ rates conditional on disclosure status, estimated using the specification in column 6 of Table 2.2. It shows that the negative effect of revealing a low wage is mostly confined to the bottom two wage deciles. This suggests to us that – while learning the $HighWage$-$LowWage$ distinction is not informative to firms about productivity– disclosing a very low wage does convey some negative productivity information to firms. While most workers seem to be aware of this and hide very low wages, the few that do disclose these wages suffer when they do.[42]

Turning next to the question of why $OverWage$ workers might choose to disclose, Section 6.2 will present evidence in support of the idea that disclosure serves a useful function for these workers despite its negative effects on success rates at any given job: It filters out offers from jobs whose wages are likely to be unacceptable to the worker, perhaps by communicating an unwillingness to accept offers below their current wages.

We report the coefficients of the other control variables in Table 2.2 in Appendix B.5.3. Overall, these estimates confirm common-sense expectations about which applications are more likely to succeed. For example, indicators of mismatch, such as being outside the employer's preferred age range, having less than the requested work experience, or being the 'wrong' gender have strong negative effects on application success rates. Also, hiring agents prefer applicants whose current location, industry and occupation match those of the job; they also prefer workers who express an interest in working in the job's location, industry or occupation, although these effects are smaller in magnitude. We also find that job seekers' investments in information acquisition –i.e consulting the board's *job lens* function– raise their success rates. Applying to jobs indiscriminately has the opposite effect: applications made through the 'batch apply' mechanism have a 0.5 percentage point lower chance of being labeled a recruiting $Target$. This effect is just as large as having less than the required experience for the job.

[42] It is also interesting to note that the negative effect of disclosure becomes insignificant for $UnderWage$ workers, but not for $OverWage$ workers, when we control for applicant fixed effects.

### 2.6.2 Wages of Successful Applications

In addition being selected as a successful candidate by a firm, a second outcome of interest to job seekers –especially those who already have a job– is the wage attached to the positions where their applications succeeded. To explore the effect of disclosure on this outcome, we now estimate the effects of disclosure on the mean wage attached to a worker's *successful* applications. If high-wage workers use disclosure to restrict the set of jobs that respond to them, disclosure might cause the mean wage of their successful applications to rise via a composition effect, even though they succeed in fewer of their applications.[43] Because each worker can accept only one job, this could be valuable to them if (contrary to our theoretical model) receiving and rejecting offers is costly; it could also be economically efficient if, in addition, issuing an offer is costly for firms.

To test this 'offer quality' hypothesis, we use the job's posted wage as a proxy for the job's quality, and define successful applications as ones achieve $Target$ or $Save$ status. We then run the following regression on the sample of successful applications:

$$\begin{aligned} Z_j =& \gamma_0 + \gamma_1 OverWage_{ij} + \gamma_2 Disclose_{ij} \times UnderWage_{ij} + \gamma_3 Disclose_{ij} \times OverWage_{ij} \\ &+ CX_{ij} + FE + e_j \end{aligned} \tag{2.3}$$

where the dependent variable, $Z_j$ is the midpoint of the posted annual wage range of job $j$ (in 10 thousands of RMB).[44] The $\gamma$ coefficients have the same interpretation as

[43] The only information we have about wages paid by the jobs workers find on Liepin is the wage range that is attached to each ad when it is created. In consequence, none of our results are informative about how wage revelation impacts wage negotiations between individual workers and employers after the employer contacts a targeted worker. Thus, in contrast to Exley et al. (2020) and Roussille (2020) who study these negotiations, our estimates of wage effects capture composition effects only: Revealing a high wage filters out recruiting interest from low-wage employers.

[44] The midpoint of posted wage is defined as (lower bound of posted wage + upper bound of posted wage)/2. We perform a robustness check which takes the lower bound and upper bound of posted job as outcome variables in Appendix B.6.1.

the $\alpha$'s in equation (2.3). Our estimates are reported in Table 2.3, which only controls for the applicant's characteristics (because the mechanism of interest is the change in the mix of jobs in which the applicant succeeds). In more detail, column 2 controls for the applicant's gender and current wage. Column 3 adds quadratics in the applicant's age and marital status. Column 4 controls for the applicant's education level, work experience and website classification variables. We add controls for the gap between the applicant's desired and current wage, the match between the applicant's current location, industry and occupation and the desired ones, and whether the application is from batch apply and whether it is made after using job lens in column 5. In column 6, we add fixed effects for the applicant's current location, industry and occupation, and the application time. Column 7 replaces the applicant's characteristics with a full set of worker fixed effects.[45] Standard errors are clustered at the worker's sub-occupation level.

Unsurprisingly (and consistent with Table 2.2), the $OverWage$ coefficients in Table 2.3 show that workers with higher current wages (relative to their co-applicants) tend to succeed in jobs that pay more, even when workers don't disclose their wages. This is true in all specifications and highly statistically significant. Specifically, among workers who chose to conceal their current wages, column 7 indicates that the successful applications of $OverWage$ workers pay between 9,899 to 13,143 RMB more per year than the successful applications of $UnderWage$ workers. This amounts to a 5.2% to 6.9% increase (the midpoint of the posted wage is 191,854). Also unsurprisingly, *revealing* a low wage reduces the mean wage of successful applications in all specifications, although these effects become mostly insignificant in the most saturated specification. Most relevant to the offer quality hypothesis, we find that –while disclosing a

[45]In column 7, the effects of wage disclosure are identified only by workers who change their disclosure status. The effects of $OverWage$ status, however, are identified from a much larger set of workers because this varies across applications, within workers, regardless of whether wages are disclosed.

high wage appears to reduce mean offered wages in the absence of controls (column 1), these estimates reverse once we control for a few basic applicant characteristics, such as education and work experience. Specifically, according to column 7, disclosing a high wage raises the mean wage of successful applications by 2,049 to 2,925 RMB, which amounts to 1.1% to 1.5% of the midpoint of the posted wage. This supports the notion that high-wage workers use wage disclosure to filter out unwanted low-wage offers.

### 2.6.3 Heterogeneity and Robustness Checks

So far Section 6, we have demonstrated two main facts describing how employers react to applicants' relative wages and wage disclosure decisions: (1) employers appear to be attracted to unexpectedly high-wage applicants (regardless of whether the applicants disclose their wages), and (2) high-wage workers appear to use wage disclosure as a filtering device that discourages low-wage employers from targeting them for recruitment.

To shed additional light on these findings, we now probe their sensitivity to two factors that affect firms' ability or willingness to pay high wages: firm size and labor market tightness. To that end, Figures B.12 - B.14 display the estimates of our three main coefficients in Table 2.2 –$\alpha_1$, $\alpha_2$, and $\alpha_3$– for different firm sizes and labor market tightness categories.

Focusing first on the *filtering* effects of *disclosing* a higher-than-expected wage ($\alpha_3$), panel (c) of Figure B.12 shows that disclosing such a wage reduces the chances of being targeted most strongly in the two smallest firm size categories in our data, i.e. among the firms who are least likely to be able t afford high wages. Similarly, panel (c) of Figures B.13 and B.14 show that this negative effect is strongest in slack labor markets (whether measured at the industry or occupation level). Employers avoid targeting workers who have a high current wage when plenty of other workers are available for hire. Together, these findings increase our confidence that some high-wage workers are using wage disclosure as a way to filter out unwanted, low-wage job offers.

Turning next to employers' attraction to applicants with unexpectedly high current wages –even among workers who choose not to reveal their wages– ($\alpha_1$), panel (a) of Figure B.12 shows that this attraction is present among firms of all sizes, but is especially *strong* in small firms: such firms may infer a willingness to be accept an offer from the fact that the worker chose not to reveal his high wage. Figures B.13 and B.14 show that employers' attraction to (non-disclosing) high-wage workers is present in both tight and slack labor markets, with some suggestion of being weaker in slack labor markets.

We next ask whether the high-wage *disclosure* penalty varies with the applicant's gender. If gender stereotypes associate women with modesty and low wages, then employers might react more negatively to disclosures of a high wage by women, compared to men. To answer this question we add the full set of interactions between $Male$, $Disclose$ and $OverWage$ to the regressions in Table 2.2 and report the results in Table B.21. In general, the effects of $OverWage$, $Disclose$ and their interactions are consistent with the baseline specifications in Table 2.2. (For example, firms like workers with unexpectedly high wages, regardless of wage disclosure status.) However, in almost all cases –and especially when we control for both worker and job fixed effects in column 7– all the interactions with gender are statistically insignificant.

## 2.7 Discussion

This paper has investigated workers' voluntary wage disclosure decisions during the job application process in a real, high-stakes labor market. Using data from Liepin.com, we develop three alternative heuristics an employer might use to impute wages to the non-disclosing workers in an applicant pool. Using these heuristics to study workers' wage revelation decisions, we find that, on average, workers behave *as if* firms interpret an unexpectedly high current wage as a signal of higher worker productivity: Workers tend to hide low wages and reveal high ones. We also find that men are substantially more likely to disclose their wages than women at all wage levels; this gender gap is hard to explain by similar sorts of strategic considerations.

Turning to firms' reactions to workers' wages and disclosure decisions, we document two surprising patterns: First, while employers *are* more likely to prefer workers with unexpectedly high wages (suggesting that wage residuals are positively correlated with worker productivity), this preference holds *equally* among workers who disclose and who hide their wages. Thus, it appears that employers can infer the idiosyncratic differences in worker productivity that are reflected in workers' current wage levels from other available information, including the workers' resumes. Second, we find that the act of disclosing one's wage *reduces* the application success rates of *both* low- and high-wage workers. We argue that the disclosure penalty for low-wage workers is most likely driven by a small number of very low-wage workers who make the mistake of disclosing their wage. The disclosure penalty among high-wage workers, however, is more consistent with a deliberate decision by these workers to screen out job offers that pay less than the worker's current wage.

As already noted, our results have a number of policy implications, including the fact that men would be disproportionately constrained by policies that prohibit or discourage voluntary wage disclosure; men however may be the *beneficiaries* of salary history bans, which prohibit firms from asking about wages but do not prevent voluntary disclosure. With respect to wage levels, our results on who is directly impacted are more subtle: recall that workers' *absolute* wage levels had only minuscule effects on salary disclosure rates. Instead, the *HighWage* workers who are most constrained by salary history bans are workers who have chosen *conservative* application strategies (i.e. who apply to jobs where they are better paid than most of the other applicants). These 'timid' workers may not be the best paid in absolute terms. Our finding that employers appear to know a worker's current wage even when it is not disclosed also suggests that –perhaps surprisingly– a worker's current salary may not be a major source of asymmetric information in labor markets. As in the proverbial used car market –which now has access to rich information about a vehicle's history– advances in information technology and market design may have made this information relatively easy for buyers to access or infer.

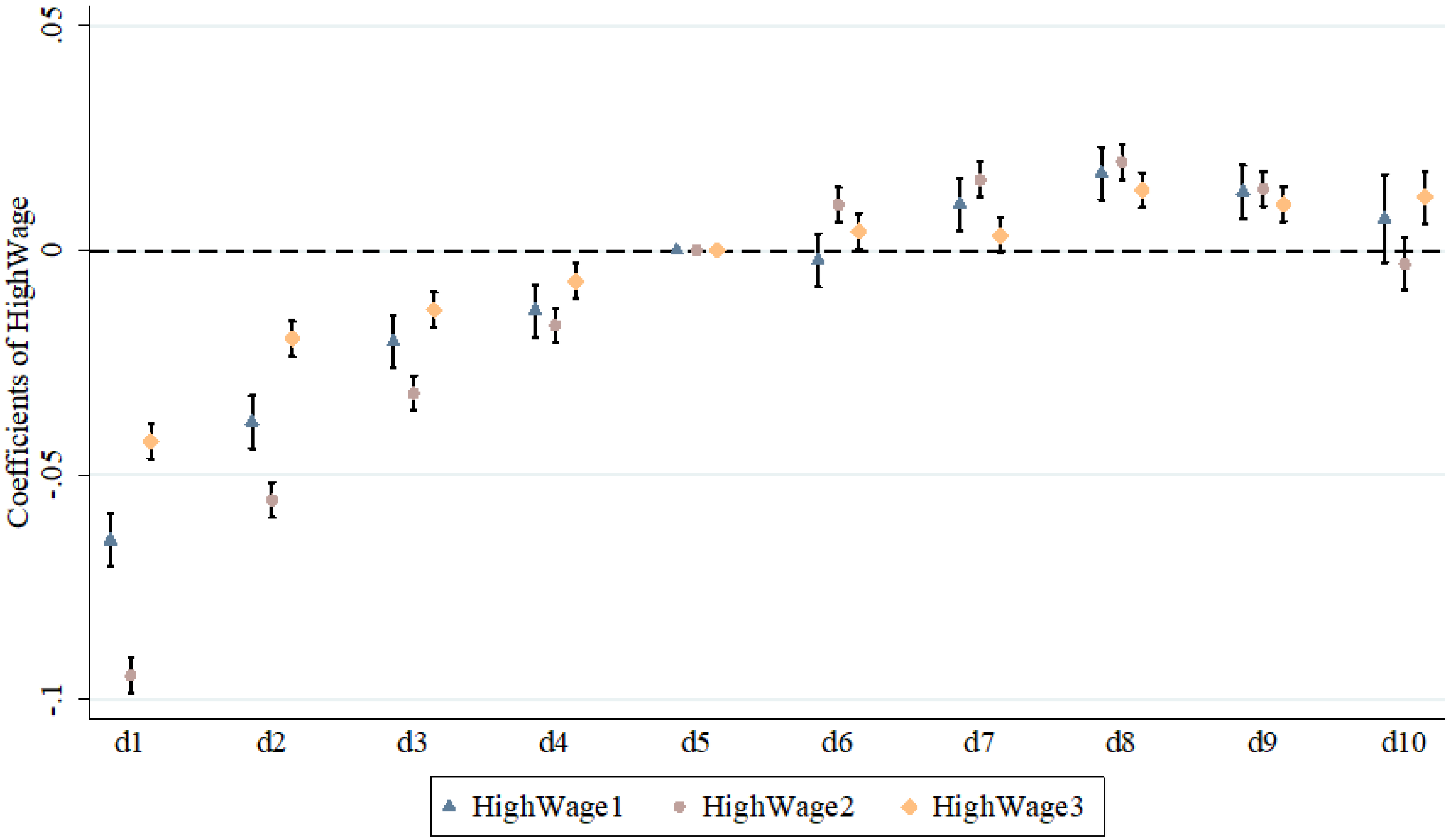
One important limitation of our analysis is the fact that we do not observe job offers or final hiring decisions, which prevents us from studying the wage bargaining process that could occur between the candidate-selection and hiring stages. In this respect our work is complementary with Roussille (2020), who focuses on this later part of the recruitment process. Similarly, while our results on the effects of wage disclosure address the same question as the resume audit study in Agan et al. (2021), our results on the determinants of workers' *disclosure* are complementary to their work, which does not study workers' decisions on what to reveal in their resumes.

While this paper has studied workers' *wage* revelation decisions and their consequences, the methods we have developed can be applied to a long list of verifiable worker characteristics whose disclosure on resumes is neither automatic nor taboo. These include union organizing activity, criminal history, pregnancy, college grades, gaps in work history, credit history, LGBTQ identity, appearance (e.g. a photo), age, marital status, responsibility for children, disability, and medical history. For example, one important lesson from our work is our demonstration that *disclosure decisions are much less sensitive to the absolute level of a characteristic than to the unexpected, or residual component of that characteristic*: for example, a credible disclosure of no criminal history can be much more helpful to members of groups with high levels of criminal activity than other groups. A second lesson is that *firm-worker matching matters*: because any one of these characteristics may be viewed positively by some employers and negatively by others, a worker's optimal revelation strategy depends on *where she intends to apply*. Finally, our analysis has demonstrated that in *workers' optimal disclosure decisions must consider two distinct consequences of disclosure: the wage offer a worker is likely to receive, and the chances of receiving an offer*. Depending on a workers' actual wage relative to the employer's expectations, these effects may not work in the same direction.

Finally, we note that that some resume characteristics, like education and work history, appear to be universally expected in resumes (and thus conspicuous when they are absent) while others (like age and marital status) are rarely seen –at least in the U.S.– and are thus conspicuous when they are present. In addition, these conventions (such as whether age and a photo are expected on a resume) vary across countries and over time. This raises the issue of how (in addition to government legislation and platform design) social conventions affect workers' disclosure decisions, and how these conventions affect employers' interpretation of a worker's decision to disclose. For example, if it is not typical for workers to disclose criminal history, how will employers treat a resume that affirmatively claims "no criminal history"? Exploring the effects of such conventions might shed additional, interesting light on the complex consequences of asymmetric iformation in labor markets.

Figure 2.1: Wage Disclosure Setting in Liepin.com

Figure 2.2: The Effect of *HighWage* Deciles on Wage Disclosure



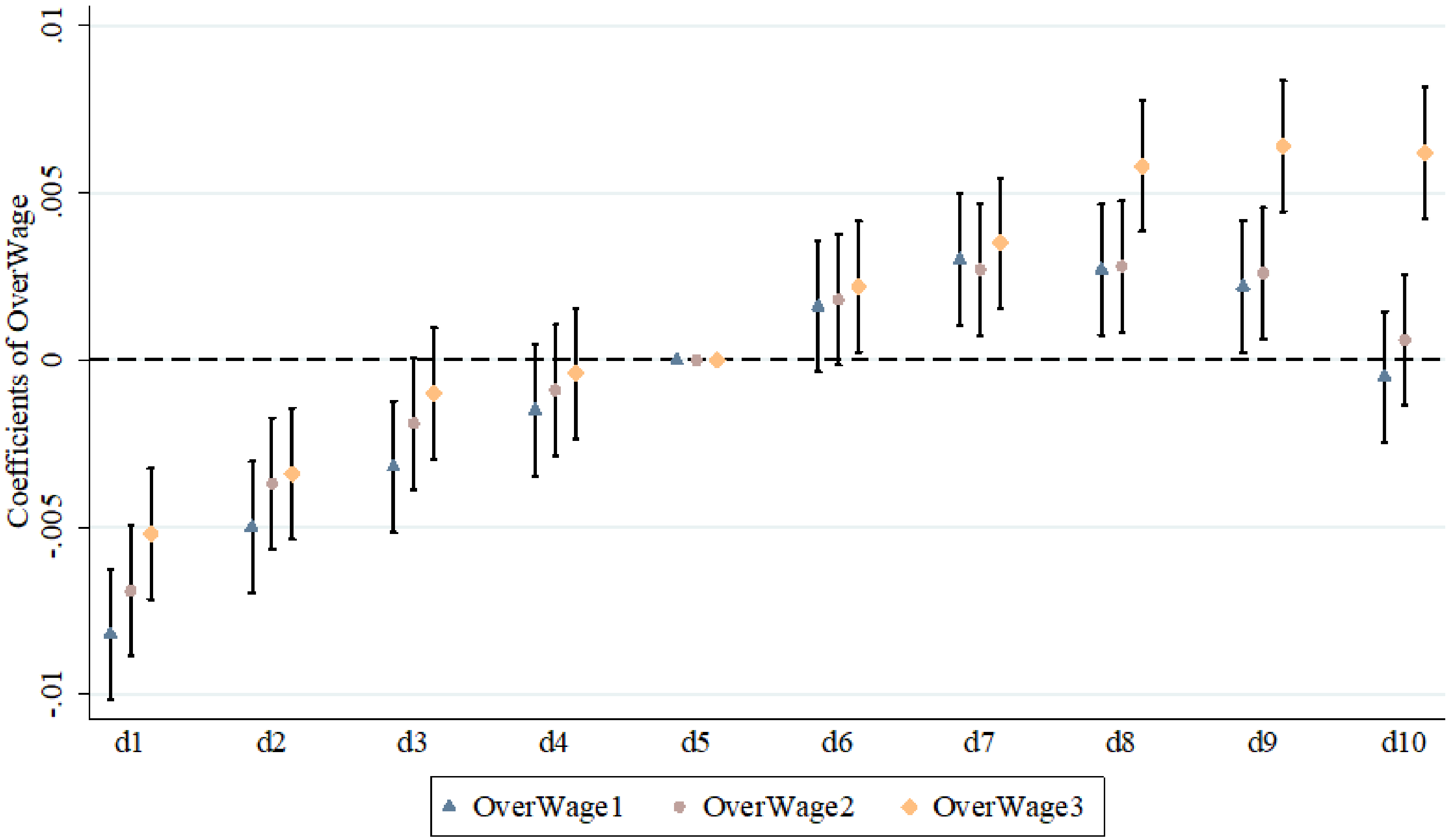

Notes:

1. Figure 2.2 plots the coefficients of HighWage deciles of the same regression in column 7 of Table 2.1, except for replacing

binary HighWage with ten deciles dummies.

2. Consistent with HighWage indicators proposed in Section 5.1, HighWage1 deciles are generated from the distribution of the lower bound of posted wages of jobs that the worker has applied for. HighWage2 deciles are generated from the distribution of wages of applicants that have applied for the same jobs. HighWage3 deciles are generated from the distribution of normalized wage prediction residuals, in which the normalized residual is defined as (actual wage – predicted wage)/predicted wage (as shown in Appendix B.4.1).

Figure 2.3: The Effect of *OverWage* Deciles on Becoming a Recruiting Target




Notes:

1. Figure 2.3 plots the coefficients of OverWage deciles of the same regression in column 6 of Table 2.1, except for replacing

binary OverWage with ten deciles dummies.

2. Consistent with OverWage indicators proposed in Section 6, OverWage1 deciles are generated from the distribution of applicants' wages with the lower posted wage as decile 5. OverWage2 deciles are generated from the distribution of wages of applicants for the job. OverWage3 deciles are generated from the distribution of normalized wage prediction residuals, in which the normalized residual is defined as (actual wage – predicted wage)/predicted wage (as shown in Appendix B.5.1).

Table 2.1: The Effect of Applicant's Characteristics on Wage Disclosure

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| | | | | **Panel A** | | | |
| HighWage1 | 0.0472*** | 0.0494*** | 0.0593*** | 0.0533*** | 0.0477*** | 0.0514*** | 0.0510*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Wage | | -0.0048*** | -0.0034*** | -0.0018*** | -0.0009*** | -0.0004*** | -0.0004*** |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Male | | 0.0826*** | 0.0942*** | 0.0734*** | 0.0912*** | 0.0843*** | 0.0884*** |
| | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| | | | | **Panel B** | | | |
| HighWage2 | 0.0288*** | 0.0311*** | 0.0400*** | 0.0421*** | 0.0405*** | 0.0417*** | 0.0450*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Wage | | -0.0046*** | -0.0033*** | -0.0018*** | -0.0009*** | -0.0004*** | -0.0004*** |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Male | | 0.0822*** | 0.0935*** | 0.0726*** | 0.0907*** | 0.0839*** | 0.0880*** |
| | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| | | | | **Panel C** | | | |
| HighWage3 | 0.0292*** | 0.0870*** | 0.0645*** | 0.0310*** | 0.0314*** | 0.0154*** | 0.0157*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Wage | | -0.0052*** | -0.0038*** | -0.0018*** | -0.0009*** | -0.0001* | -0.0001 |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Male | | 0.0880*** | 0.0949*** | 0.0736*** | 0.0909*** | 0.0837*** | 0.0877*** |
| | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Demographics | | | Yes | Yes | Yes | Yes | Yes |
| Edu & Exp | | | | Yes | Yes | Yes | Yes |
| Classification & Match | | | | | Yes | Yes | Yes |
| Location FE | | | | | | Yes | Yes |
| Industry FE | | | | | | | Yes |
| Occupation FE | | | | | | | Yes |
| 'Effective' *N* | 941,733 | 941,733 | 941,733 | 941,733 | 941,733 | 941,697 | 941,695 |

Standard errors in parentheses, clustered by worker's sub-occupation. *** p<0.01, ** p<0.05, * p<0.1

Notes:

1. HighWage1 represents workers whose wages are above the minimum posted wage of the median job in their application sets. HighWage2 represents workers whose wages are greater than the median wage of applicants who have applied the same jobs. Highwage3 represents workers whose wages are higher than the predicted wage based on characteristics in their resumes.

2. In Table 2.1, column 3 includes worker's marital status and a quadratic in age. Column 4 adds education and experience variables, including the highest degree, whether the highest degree is Tongzhao, domestic and world rank of the university that the applicant achieved her degree, and whether the university is 985/211; employment status,

years of working experience, tenure and industry in the last two jobs. Column 5 includes web classification and match variables: whether the resume is elite, how long the resume has been created, profile completeness score, if the job seeker has a golden membership, the number of applications, the gap between the desired wage and current wage, and match variables that measure the alignment between the applicant's current location/industry/occupation and the desired ones. Fixed effect of worker's location is added in column 6, and fixed effects of worker's industry and occupation are included in column 7.

3. See Appendix B.2.2 for the complete version of Table 2.1.

Table 2.2: The Effect of Wage Disclosure on Becoming a Recruiting Target

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| | | | | **Panel A** | | | |
| OverWage1 | 0.0098*** | 0.0078*** | 0.0067*** | 0.0078*** | 0.0057*** | 0.0059*** | 0.0034*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.000) | (0.000) | (0.000) |
| Disclose* UnderWage1 | -0.0033*** | -0.0046*** | -0.0029*** | -0.0022*** | -0.0029*** | -0.0029*** | -0.0011 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.001) |
| Disclose* OverWage1 | -0.0029*** | -0.0043*** | -0.0030*** | -0.0025*** | -0.0029*** | -0.0031*** | -0.0023** |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.001) |
| | | | | **Panel B** | | | |
| OverWage2 | 0.0057*** | 0.0054*** | 0.0042*** | 0.0054*** | 0.0046*** | 0.0050*** | 0.0024*** |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Disclose* UnderWage2 | -0.0032*** | -0.0047*** | -0.0029*** | -0.0024*** | -0.0029*** | -0.0028*** | -0.0011 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.001) |
| Disclose* OverWage2 | -0.0028*** | -0.0043*** | -0.0030*** | -0.0023*** | -0.0029*** | -0.0031*** | -0.0022** |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.001) |
| | | | | **Panel C** | | | |
| OverWage3 | 0.0073*** | 0.0072*** | 0.0067*** | 0.0069*** | 0.0066*** | 0.0069*** | 0.0019*** |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Disclose* UnderWage3 | -0.0037*** | -0.0048*** | -0.0031*** | -0.0023*** | -0.0027*** | -0.0025*** | -0.0014 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.001) |
| Disclose* OverWage3 | -0.0035*** | -0.0047*** | -0.0032*** | -0.0025*** | -0.0032*** | -0.0034*** | -0.0025** |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.001) |
| Job Characteristics | | Yes | Yes | Yes | Yes | | |
| Job Match | | | Yes | Yes | Yes | Yes | Yes |
| Worker Characteristics | | | | Yes | Yes | Yes | |
| City, Time, Ind & Occ FE | | | | | Yes | | |
| Firm FE | | | | | Yes | | |
| Job FE | | | | | | Yes | Yes |
| Worker FE | | | | | | | Yes |
| 'Effective' *N* | 3,542,049 | 3,542,049 | 3,542,049 | 3,542,049 | 3,541,342 | 3,486,460 | 3,201,726 |

Standard errors in parentheses, clustered by job. *** p<0.01, ** p<0.05, * p<0.1

Notes:

1. Overwage1 is an indicator for applicants with wage greater than the job's lower bound of posted wage. Overwage2 represents applicants with wage above the median wage of applicants to the job. Overwage3 indicates applicants with wage greater than the fitted wage on job level.

2. In Table 2.2, column 2 controls for applicant's gender and job's characteristics including requirements for age, education, and working experience; the offered wage range and whether the wage is visible to applicants, the number of position's subordinates and the reported feedback days. Column 3 controls for the match indicators between the applicant and the job: whether the applicant satisfies the job's gender, education, age and experience requirements, and whether her current and desired location (industry, occupation) are consistent with the location (industry, occupation) of the job. Column 4 adds

variables for applicant's other characteristics including marital status, employment status, industry and tenure of the last two jobs, education quality, the website classification variables, and controls for batch apply and the usage of job lens. Column 5 adds job's location, industry, occupation and firm fixed effects, and the fixed effect for the date of application. Column 6 replaces job's characteristics and location, industry, occupation and firm fixed effects with job fixed effect. In column 7, we drop worker's characteristics and include fixed effects for time, job and worker.

3. See Appendix B.5.3 for the complete version of Table 2.2.

Table 2.3: The Effect of Wage Disclosure on the Posted Wage in Successful Applications

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| | | | | **Panel A** | | | |
| OverWage1 | 0.7536*** | 1.5278*** | 1.5319*** | 1.5135*** | 1.5102*** | 1.4752*** | 1.3143*** |
| | (0.186) | (0.224) | (0.220) | (0.218) | (0.218) | (0.216) | (0.196) |
| Disclose* UnderWage1 | -3.9539*** | -2.8452*** | -2.4977*** | -1.9087*** | -1.8822*** | -1.5025*** | -0.4215 |
| | (0.120) | (0.083) | (0.082) | (0.079) | (0.079) | (0.077) | (0.261) |
| Disclose* OverWage1 | -2.7956*** | 0.3226*** | 0.5693*** | 0.4939*** | 0.4988*** | 0.3249*** | 0.2925** |
| | (0.058) | (0.054) | (0.054) | (0.053) | (0.053) | (0.052) | (0.136) |
| | | | | **Panel B** | | | |
| OverWage2 | 0.3553*** | 0.9597*** | 0.9630*** | 0.9587*** | 0.9559*** | 0.9284*** | 0.9899*** |
| | (0.077) | (0.177) | (0.171) | (0.168) | (0.167) | (0.162) | (0.159) |
| Disclose* UnderWage2 | -3.8299*** | -3.0240*** | -2.6766*** | -1.9033*** | -1.8682*** | -1.4853*** | -0.4565* |
| | (0.096) | (0.074) | (0.072) | (0.067) | (0.067) | (0.065) | (0.268) |
| Disclose* OverWage2 | -3.1518*** | -0.2753*** | -0.0597 | 0.4107*** | 0.4194*** | 0.4522*** | 0.2914*** |
| | (0.082) | (0.065) | (0.064) | (0.065) | (0.064) | (0.063) | (0.101) |
| | | | | **Panel C** | | | |
| OverWage3 | 0.3806*** | 1.9608*** | 1.5366*** | 1.3689*** | 1.3362*** | 1.8679*** | 1.2588*** |
| | (0.082) | (0.160) | (0.154) | (0.149) | (0.148) | (0.147) | (0.146) |
| Disclose* UnderWage3 | -3.4499*** | -2.4938*** | -2.3438*** | -1.7302*** | -1.6982*** | -1.3754*** | -0.1924 |
| | (0.098) | (0.075) | (0.074) | (0.070) | (0.071) | (0.068) | (0.279) |
| Disclose* OverWage3 | -3.1995*** | -0.1313* | -0.0225 | 0.3682*** | 0.3664*** | 0.3649*** | 0.2049** |
| | (0.094) | (0.073) | (0.072) | (0.071) | (0.071) | (0.070) | (0.100) |
| Wage & Gender | | Yes | Yes | Yes | Yes | Yes | |
| Demographics | | | Yes | Yes | Yes | Yes | |
| Edu, Exp, Classification | | | | Yes | Yes | Yes | |
| Match & Application | | | | | Yes | Yes | |
| City, Time, Ind & Occ FE | | | | | | Yes | |
| Worker FE | | | | | | | Yes |
| 'Effective' *N* | 417,723 | 417,723 | 417,723 | 417,723 | 417,723 | 417,676 | 266,713 |

Standard errors in parentheses, clustered by worker's sub-occupation. *** p<0.01, ** p<0.05, * p<0.1

Notes:

1. In Table 2.3, column 2 controls for the applicant's gender and current wage. Column 3 adds controls for the applicant's quadratic in age and marital status. Column 4 controls for the applicant's education level, work experience and website classification variables. Column 5 controls for the gap between the applicant's desired and current wage, the match between the applicant's current location, industry and occupation and the desired ones, and whether the application is from batch apply and whether it is made after using job lens. Column 6 adds fixed effects for the applicant's current location, industry and occupation, and the application time. Column 7 replaces the applicant's characteristics with worker fixed effect.

2. The effects of wage disclosure on the lower bound and upper bound of posted wage in successful applications are examined in Appendix B.6.1.

# Chapter 3

# Measuring Algorithmic Bias in Job Recommender Systems: An Audit Study Approach

## 3.1 Introduction

With the rapid development of the Internet, the explosive growth of information makes it increasingly challenging for people to process a huge amount of data and to find desired information, products and workers. The personalized recommender system, first proposed in the 1990s, is a powerful tool to alleviate the information overload problem by prioritizing the delivery of information and showing every user a different list of new items that match her personal interests and preferences (Lee and Brusilovsky, 2007). Recommender systems have been widely and successfully applied in online websites and e-commerce services. For instance, a customer on Amazon possibly sees a page called "Customers Who Bought This Item Also Bought," which displays the products that she is likely to be interested in. After people watched a movie on Netflix, it often suggests people what to watch later, called "People Who Liked This Movie Also Saw"(Jannach et al., 2010).[1]

Similar scenarios can be found on internet-based recruiting platforms, which have now accumulated a vast volume of information on workers and jobs. According to statistics from Glassdoor.com, in the US, there were 2.09 million jobs posted online by employers in 2019, and more than half of job seekers preferred finding job opportunities on online job sites.[2] In addition, the wide usage of online job searching and recruiting enables internet job boards to characterize behaviors and activities of job seekers and employers, which together foster the development of job recommender systems. Job recommender systems apply the concept of personalized recommendation to the job recruiting domain to suggest better matches between job seekers who search for job positions and recruiters who find candidates on the Internet. Virtually all internet job boards now recommend jobs to the workers who use their platforms. These customized recommendations are generated by algorithms, using criteria that include the worker's characteristics and previous behaviors, and the match between the worker's characteristics and the job's requirements. While job recommendation algorithms have the potential to help workers and firms find better matches faster, they also have sparked deep concerns about fairness: even when there is no discriminatory intent from designers, the recommended jobs may reinforce gender and other stereotypes. For instance, in content-based recommendation algorithms, gender might be associated with certain types of jobs and specific personalities in the workplace, which

[1] Recent evidence shows that 35% of purchase on Amazon and 80% of stream time on Netflix are driven by the recommendation systems. See https://towardsdatascience.com/deep-dive-into-netflixs-recommender-system-341806ae3b48 and https://www.mckinsey.com/industries/retail/our-insights/how-retailers-can-keep-up-with-consumers.

[2] See Glassdoor's HR and Recruiting Stats for 2020 https://www.glassdoor.com/employers/resources/hr-and-recruiting-stats/reasons-to-use-glassdoor, and Glassdoor's Job & Hiring Trends for 2020 https://www.glassdoor.com/research/app/uploads/sites/2/2019/11/Job_Hiring_Trends_2020-FINAL-1-1.pdf.

leads to gender segregation in job recommendations (Chaturvedi et al., 2021; Gaucher et al., 2011). Furthermore, based on job seekers' application behaviors, item-based collaborative filtering algorithms, as well as algorithms that incorporate the past behaviors of hiring agents, can create and perpetuate previous gender differences in recommendations received by workers.

This paper measures whether, to what extent, and how job board algorithms systematically treat male and female job seekers differently by conducting an *algorithm audit*, which is a new research approach proposed in recent years to study the black-box of algorithm features and to ascertain whether algorithms result in harmful discrimination by using fictitious correspondence in online platforms (Sandvig et al., 2014; Hannák et al., 2017). More specifically, I created otherwise identical male and female worker profiles on the four largest Chinese job boards, and observed which jobs were recommended to those profiles. In each job board, I selected 35 types of jobs based on three criteria: the number of active job openings, the job's gender-type (female-dominated jobs, gender-balanced jobs and male-dominated jobs), and hierarchy level (entry, middle, and high). Then I created resumes that were qualified for the above jobs; these come in pairs that are identical except for applicant gender. Since Chinese employers' gender preferences appear to interact strongly with the worker's age (Helleseter et al., 2020), I made two versions of each profile pair — a 'young' version and an 'older' version, in which the older applicants have 10 more years of working experience than young applicants. In order to track how algorithms update their recommendations based on workers' application behaviors, my fictitious workers then applied for the top jobs in their recommendation lists. I repeated this application process up to three times (each time responding to a new set of recommendations), then compared the job recommendations received by male and female applicants.

I find that identical male and female applicants do not always receive the same job recommendations: out of 100 job recommendations received by my applicants, 9.72 jobs were uniquely displayed to male or female applicants. Senior workers, who have more years of working experience, received a smaller number of gender-specific recommendations. Importantly, gender divisions in recommendations are even higher after fictitious applicants started applying for jobs: The raw difference rate between male and female applicants is 7.7% in the first round, whereas after three rounds of applications, 18.4 percent of recommendations are gender-specific. Because jobs displayed at the top of the recommendation list receive more attention, I further define the *list difference* in job recommendations, in which two job recommendations are the same only if both the job and the rank are identical in the recommended lists for pairwise workers (i.e., the third job in the men's list is the same with the third job in the women' list), and find that around three in four recommendations are different across male and female applicants.

To detect gender bias in the quality of recommended jobs, I leverage statistical tests to quantify the gender gap of both explicit and implicit measures of job quality. *Explicit measures* include the job's posted wage, requested education, and requested working experience. I find that on average, only-to-male jobs, which are seen by men rather than women, posted wages that were 1.9% higher than jobs recommended to women; this difference is marginally statistically significant. While the requested education is the same in jobs recommended to male and female applicants, jobs recommended only to men have 0.08 more years of working experience requirement than only-to-female jobs.

Furthermore, since job descriptions implicitly convey information on job quality, I extracted words used in the job descriptions reflecting five aspects pf quality: *skills*, *benefits*, *work form*, *company information*, and *other requirements*. By comparing the word frequency in male-only and female-only job ads, I find that literacy skills and administrative tasks are more likely to show up in female-only jobs, while influencing skills such as leadership and decision-making are mentioned more in male-only jobs. On the other hand, female applicants are recommended to apply for more jobs with flexible working hours and normal breaks in comparison to men with identical characteristics, while male applicants see more jobs that need night work and overtime. For benefits, only-to-female jobs place more emphasis on base pay, marriage leave, and parental leave, while only-to-male jobs focus on more performance incentives such as reward and company stocks or options. Company-related words do not significantly differ between male-only and female-only jobs, except that orientation training is involved in more female-only jobs, while male-only jobs are more likely to be in publicly-listed companies.

The other requirements contained in the job descriptions also reflect gender-based differences in job recommendations. Words in jobs recommended to women are often related to feminine personality, such as *patient* and *careful*, and have more descriptions on desired workers' appearance such as *facial features*, *figure*, and *temperament*. Jobs recommended to men prefer workers who are *self-motivated*, *experienced*, and are able to *work under pressure*. Moreover, these male and female words in recommended jobs are consistent with gendered words summarized in previous literature in language (Fitzpatrick et al., 1995), in political science (Roberts and Utych, 2020), in psychology (Rudman and Kilianski, 2000) and in labor economics (Gaucher et al., 2011; Kuhn et al., 2020; Chaturvedi et al., 2021). To collect the gendered perceptions of words, I conducted two surveys on Amazon MTurk and on Chinese workers, and found that feminine words emerge more within jobs seen by female applicants and jobs recommended to men contain more masculine words. This suggests that words used in gender-specific jobs are associated with widely held gender stereotypes in the workplace, and the inclusion of stereotype-linked words contributes to the gender bias in job recommendation systems.

Finally, I attempt to isolate the precise mechanisms accounting for gender bias in job recommendations. *Content-based recommendations*, which link gender with jobs' features must play a role because words about gender-related personality traits (i.e., patient in female, work under pressure in male) and gender stereotypes in the workplace (i.e., women are good at literacy skills, men have leadership) occur differently in gender-specific recommendations. Moreover, hiring agents' behaviors also appear to contribute to gender-biased job recommendations. When more hiring agents read their profiles, the pairwise male and female applicants will see more different job ads in their recommendations, indicating that human bias may be maintained in and interact with recommender systems. Lastly, by comparing jobs recommended before and after workers apply for jobs, I find that *item-based collaborative filtering* which recommends jobs based on workers' application history may reinforce and amplify the gender bias in the system.

This paper is related to four existing literatures. The first is the broad literature about gender inequality in labor markets. Using both traditional survey data and internet job board data, this literature has documented that gender inequality is accentuated by gender differentials in job search patterns, such that women are less likely to search for jobs outside of their living places and switch occupations (Eriksson and Lagerström, 2012), and women have higher levels of risk aversion in accepting offers (Cortés et al., 2021), from gender discrimination in the recruiting process in which employers prefer men in some certain occupations (Booth and Leigh, 2010; Cediey and Foroni, 2008), from gender segregation in skills (Christl and Köppl-Turyna, 2020; Stinebrickner et al., 2018), from gender differences in workplace bargaining propensity (Card et al., 2016), and from family burdens in promotions and career development (Petit, 2007). As far as I know, this is the first paper to study gender bias in job recommendations. While existing literature studies gender differentials at various stages of the search and matching process, I argue that gender differences and gender discrimination can arise even at the very early stage, where male and female workers may see different job vacancies in online job platforms due to the personalized job recommendations. More importantly, when the algorithm predicts workers' preferences based on their previous behaviors, feedback loops and self-fulfilling prophecies in recommendation algorithms may magnify the gender bias (Cowgill, 2018; Jiang et al., 2019), in which gender differences in job applications can yield to greater gender bias in the future job recommendations.

Methodologically, my paper contributes to the audit studies (or correspondence studies), which are widely used in the research on discrimination in social sciences. Aiming at comparing callback rates from real employers between two identities, audit studies have to create resumes that are as close as possible to real workers, and the detailed information, such as working experience on resume, is always randomly selected from resume banks (Gaddis, 2018). Due to the complexity of resume design and the high cost of callback collection, most audit studies only focus on a few occupations and industries, especially entry-level and unskilled jobs in manufacturing and service sectors; therefore, evidence on gender discrimination is lacking for senior-level and high skilled jobs which require proof of identity or qualifications (Rich, 2014). Compared to previous audit studies, my algorithm audit has three advantages: First, my resume design is much easier as the fictitious resumes only include the minimum information that is required by job platforms rather than any detailed descriptions of workers' personal working histories and statements. Second, since my goal is to investigate job recommendation outcomes from the workers' side, this study has no contacts with employers and does not collect callbacks from employers, which avoids alerting employers to the experiment (Avivi et al., 2021). Finally, the field experiment was performed on the four largest Chinese job boards and chose 35 job types in each platform, ranging from unskilled jobs such as sales and warehouse keeper, to high-level jobs such as financial manager and software engineer, which covered a broad and representative sample in online labor markets in China.

In addition, this paper contributes to the emerging literature on algorithmic fairness in economics. With the increasing engagement of algorithms in supporting human decision making, algorithmic bias and fairness have been studied in various fields such as advertisement delivery (Lambrecht and Tucker, 2019), criminal courts (Angwin et al., 2016) and mortgage approval (Fuster et al., 2020; Bartlett et al., 2021). In labor markets, existing research mainly focuses on gender bias in algorithms used in recruitment and in performance evaluation. For instance, Li et al. (2020) develop a resume screening algorithm that explicitly values exploration and show that efficiency (the quality of interview decisions) and equity ( demographic diversity of applicants) can be improved at the same time in the workplace. Prassl (2018) documents that the evaluation algorithms in Uber result in lower payments for female drivers. However, to my knowledge, there is no research about the fairness of job recommendation algorithms from the perspective of job platforms. My research fills this gap by demonstrating that gender bias exists in the job recommendation algorithms, which comes even before workers apply for jobs. Moreover, when the hiring agents' behaviors are incorporated into the job recommender, gender discrimination in recruitment and gender bias in job recommendation interplays with each other, which potentially perpetuates gender inequality in the matching in labor markets.

Lastly, my work complements and extends research on gender equality (Poutanen and Kovalainen, 2017; Barzilay and Ben-David, 2016; Athreya, 2021; Cook et al., 2021) and algorithm transparency (Tambe et al., 2019; Kellogg et al., 2020) in the platform economy. From a practical point of view, few platforms in two-sided markets directly use information about gender, race or ethnicity in their algorithms. In other words, algorithmic bias is caused inadvertently in most cases. My empirical evidence from a strictly controlled experiment has important implications for platforms and policy-makers to raise their awareness of the potential dangers of systematic bias in the algorithms.

The rest of this paper is organized as follows: Section 2 discusses the related literature. Section 3 provides an overview of how the job recommender systems generate job recommendations to job seekers. In section 4, I present the potential mechanisms of the gender-biased job recommendations in online job boards. Section 5 details the experiment design and implementation. Section 6 summarizes the experimental results on the differences in job recommendations between male and female applicants. I explore the potential drivers of algorithmic gender bias in job recommendations in section 7. Section 8 concludes.

## 3.2 Literature

### 3.2.1 Gender Discrimination and Audit Studies

Audit studies, also known as correspondence studies or correspondence experiments, have been widely used to estimate discrimination on various grounds, such as race, gender and age (Fix et al., 1993).[3] In recent audit studies on gender discrimination in labor markets, researchers create fictitious workers that are identical in all dimensions except for gender and send out their resumes to real job vacancies, then any difference between male and female job candidates on the subsequent callbacks

[3]More specifically, audit studies rely on real auditors who are matched in observable characteristics, while correspondence studies create and send fictitious applications with identical variables (Bertrand and Duflo, 2017).

from employers can be interpreted as causal evidence of gender bias or discrimination (Gaddis, 2018; Baert, 2018).

Empirical evidence on gender discrimination under the framework of audit study is mixed with respect to occupation, skill level, and age. An early work from Riach and Rich (2006) used pairs of matched, written applications to test for gender discrimination in London and showed that men had fewer callbacks in female occupations, and significant discrimination against females was found in male-dominated occupations. Similar results come from Booth and Leigh (2010), suggesting that the pro-female bias exists in the occupations where the percentage of females is 80% or more in Australia, Albert et al. (2011) documenting that females are significantly preferred in lower-level, female-dominated jobs in Madrid, and Carlsson (2011) showing that women have a larger advantage in female jobs than the advantage of male in male-dominated jobs in Sweden.

Moreover, Baert et al. (2017) show that when applying for jobs at a higher occupational level, the invitations for job interviews for female applicants are about two-thirds of that their male counterparts can receive in business-related jobs in Belgium. Using the three largest Chinese job websites, Zhou et al. (2013) describe the gender discrimination heterogeneity across firms: State-owned firms prefer male applicants due to leadership, while foreign firms, firms offering marketing positions, and short-lived private firms tend to interview more female applicants.

By conducting a correspondence study in France, Petit (2007) investigates the relations between family constraint and gender discrimination in hiring. It suggests a 20% gender gap in access to job interviews in which young, single female applicants (aged 25) are less favored in high skilled administrative positions, especially in jobs offering long-term contracts, but the discrimination is eliminated in prime-age applicants (aged 37) with children. In addition, being pregnant has a substantially negative effect on the probability of being interviewed in Belgium (Capéau et al., 2012), and mothers are penalized by a lower callback rate compared to childless women and fathers in the United States (Correll et al., 2007).

### 3.2.2 Gender and Internet Job Boards

While there is plenty of literature on gender differentials and gender discrimination in labor economics (Parsons, 1991; Keith and McWilliams, 1999), the expansion of online job platforms opens up new research topics and accumulates rich sources of data on job seekers and recruiters. With respect to the worker's side, internet job boards can follow the behaviors of job seekers through the whole searching process, which allows researchers to observe and compare the labor market participation behaviors of men and women. Although there is no conclusive evidence on the gender difference in job search intensity, research built on data from online job boards has documented that women are more selective and restrictive in their choice of search area (Eriksson and Lagerström, 2012), comply more to the minimum required experience, are less open to occupational moves (Banfi et al., 2019), and are less likely to search for long duration (Faberman and Kudlyak, 2019). Moreover, results from field experiments conducted on online job platforms demonstrate the gender difference in competition and job-entry choices. Flory et al. (2015) find that women are less likely to apply for jobs with competitive compensation structure and greater earnings uncertainty. Gee (2019) find that women are more likely to finish the job application when the number of received job applicants is shown in the corresponding job posting in LinkedIn.

On the employers' side, the recruiting process that is recorded by online job boards can be divided into two phases: the attraction phase and the selection phase (Färber et al. ,2003). Attraction phase mainly refers to job posting behaviors, in which employers specify job characteristics in job ads to attract qualified employees. Kuhn and Shen (2013) studied the gendered jobs in China, which explicitly listed the gender preference in job advertisements, to examine gender discrimination. They find that men and women are equally preferred in gendered jobs, but the preference for females to males often links to youth, height, and beauty rather than offered wages and skills. In the followed studies, Helleseter et al. (2020) documented the age twist in employers' gender requests, in which gender preference shifts away from women towards men as the target age of worker rises. When employers select suitable job candidates from applicants' pool, most of the employers make callbacks to applicants with requested gender, and the gender mismatch penalty is greater for women than men (Kuhn et al., 2020). In particular, after removing the gender label in job ads, the application rate and the success rate of jobs that requested opposite gender increases for both men and women (Kuhn and Shen, 2021).

### 3.2.3 Algorithmic Fairness

Algorithmic decision-making is increasingly engaged in social and economic life, and the question of algorithmic fairness attracts plenty of research from computer science and social science. For instance, the application of algorithm tools may lead to racial bias against black defendants (Angwin et al., 2016; Cowgill, 2018), racial/ethnic discrimination in mortgage, lending and credit approval (Bartlett et al., 2021; Fuster et al., 2020), racial discrimination in health system (Obermeyer et al., 2019), algorithmic unfairness in opioid use (Kilby, 2021) and gender disparity in image search and face recognition (Kay et al., 2015; Klare et al., 2012).

Interestingly, most of the unfairness is not intended by the algorithm designers. One of the main reasons for the bias is the input data, which can be biased or unrepresentative (Kim, 2017). If the algorithm is trained on data produced by biased human decision-makers, it will reflect the bias and probably deliver bias results, as the saying goes, Bias in, bias out (Rambachan and Roth, 2019). When the characteristics for some certain groups are missing or underrepresented in training data, the algorithm's prediction on these groups is likely to be inaccurate or biased (Barocas and Selbst, 2016). In addition, interactions between users, and interactions between users and platform can also contribute to the biased results(Jiang et al., 2016).

Recently, there is growing literature about the fairness of algorithms applied in hiring. One line focuses on the adoption of algorithmic decision tools in employee selection, such as resume screening, AI interviews, evaluation on interview performance, and productivity prediction (Mann and O'Neil, 2016; Lee and Baykal, 2017; Chalfin et al., 2016; Tambe et al., 2019; Li et al., 2020). However, the resume screening tool developed by Amazon was criticized for its higher ratings for male candidates than females, which resulted from the biased training data in which Amazon hired more male workers in the past.[4] Based on the investigation on 18 vendors of algorithmic pre-employment assessments (i.e., questions, video interview analysis, and gameplay), Raghavan et al. (2020) found that most of the vendors made abstract references to "bias", but few of them explicitly revealed how to validate their models and how to fix the bias in practice.

The other line is about employers' reliance on internet platforms, and the closest work to this paper comes from Lambrecht and Tucker (2019), who conducted a field experiment on Facebook to test how online advertising algorithm delivers STEM job opportunities differently to men and women. They ran advertising campaigns targeting both men and women with otherwise identical backgrounds and found that

[4]Jeffrey Dastin, Amazon scraps secret AI recruiting tool that showed bias against women, https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G

ad about job opportunities and training in STEM was shown to 20% more men than women. While the algorithm is intended to be gender-neutral, it creates gender-biased results as a consequence of optimization cost-effectiveness in ad delivery. In advertising auction, female eyeballs have more bidders and a price premium has to be paid to show ads to women relative to men, so the STEM ads were crowded out by other advertisers in the competition. Similar results are replicated by Ali et al. (2019), who show that ads can be delivered to vastly different racial and gender audiences when Facebook optimizes for clicks. For instance, with identical ad target options, jobs in the lumber industry were delivered to an audience that was 72% white and 90% male, jobs from taxi companies reached 75% Black users, and ads for cashier positions were shown to the audience of 85% female. While the two studies focus on gender inequality under a framework of price auction in which job opportunities compete with consumer goods in commercial advertisement delivery, I study gender bias in job recommendation algorithms on internet job boards, which are the dominant platforms that match workers to jobs. Another relevant work comes from Chen et al. (2018), who explore gender equality in ranking algorithms in resume search engines using data on 855K job candidates from Indeed, Monster, and CareerBuilder based on 35 job titles in 20 U.S. cities. They find that there is a slight penalty against feminine candidates even after controlling for all other visible candidate characteristics in resumes. On the group level, the unfairness significantly benefits men in 12 out of 35 job titles. In the setting where employers proactively search for workers, ranking algorithms affect job seekers' opportunities when employers contact more the top-ranked workers (clicked into their resumes), but my paper probes into the gender inequality problem in the aspect of job seekers' applications and argue that multiple channels, more than employers' behaviors, can contribute to the gender differences in job recommendations (in terms of recommended jobs as well as their ranks).

## 3.3 An Introduction to Job Recommender Systems

Before summarizing the job recommendation algorithms, I first describe the setting in which the job recommender systems work. On internet job platforms, when a job seeker with a complete profile logs into her account, the website displays a list of jobs that the job seeker may be interested in on her homepage. Unlike the search function that requires job seekers to input keywords in the search bar, job recommendation systems generate and present recommendation results proactively and automatically.[5] Based on my personal experience with several job boards and the academic surveys on job recommender systems from Al-Otaibi and Ykhlef (2012), Hong et al. (2013) and Siting et al. (2012), most of the online job platforms build hybrid recommender systems that incorporate multiple methods.

The core and foundation of most recommender systems is *item-based collaborative filtering* (item-based CF) method. Item-based CF uses the implicit collaborations of users or items to predict the users' preferences and filters the items that are most likely of interest to users; which can be expressed as "Users who liked this item also liked". The main idea of item-based CF is to recommend items to users that are similar to the ones that the users liked in the past, where the similarity between items is derived from users' rating behaviors (Jannach et al., 2016). In the context of job recommendations, users' rating behaviors reveal how the job seeker likes a certain job. Rating behaviors, such as clicking into a job, viewing the job page, marking the job as favorite, sending a message, indicate how the worker likes a job, and of course applying for a job indicates the job seeker's strongest preference for that job. If two jobs X, Y are applied (liked) by the same job seeker, they should share some features that attract workers to apply

[5]Job recommendations play an increasingly important role in online job boards because Boolean search methods only adopt keywords to generate the results, which is insufficient and may fail to generate an appropriate match for workers and jobs (Lang et al., 2011).

both of them, so the two jobs can be defined as "similar", in other words, two jobs are "similar" if they have enough overlapped applicants. To sum up, item-based CF recommends the jobs that are similar to ones that the target job seeker applied to in the past to the job seeker.

Pure item-based CF does not function well when user's behavior data is unavailable or very sparse (e.g. newly registered job seekers and newly posted jobs).[6] To deal with the cold-start problem, most job recommender systems use a *content-based* algorithm as an important supplement to item-based CF. Based on text analysis and natural language processing techniques, content-based recommendation algorithms identify similarity between two documents by comparing the keywords in the documents, in which "content" refers to the descriptions of items' characteristics and the users' profiles.[7] In online job boards, content similarity can be established between jobs, between workers, and between jobs and workers. Two jobs are defined as similar when the same keywords appear in the job descriptions. If a job seeker applies to one of the jobs, similar ones will be recommended to him because he should have a consistent preference on jobs' content. Two workers are similar if their resumes have the same keywords, and the jobs one job seeker applies to will be recommended to the other one, since two job seekers with similar resumes should share similar tastes. Moreover, the content-

[6] If no user information is available, for instance, browsing as guests or newly registered job seekers, the knowledge-based recommendation will be used to list jobs that satisfy the user's requirements on jobs, such as the job's location, wage, and occupation. For newly posted jobs, it uses two methods to overcome the ramp-up problem. One is to apply content-based recommendations to find old jobs that are similar to the new one and recommends the new job to applicants who have already applied to these old jobs. The second is to rely on cooperation with the search algorithm. The search algorithm gives more weights for newly posted jobs to encourage job seekers to apply for new jobs when they contain certain keywords of the request of job seekers. After gathering some initial ratings, these new jobs will enter into item-based CF and can be recommended to other job seekers.

[7] Some researchers frame content-based recommendations as a classification problem of the user's likes and dislikes, and the goal is to find the classifier based on item characteristics. In this line, lots of supervised machine learning techniques such as Bayesian Classifiers, clustering, decision trees, and artificial neural networks can be applied to train models which can automatically decide whether a user is going to like a certain item.

based method also utilizes *job-worker match* attributes to make recommendations. For example, if a job ad and the worker's resume contain the same keyword, such as a skill, then the system will suggest the job seeker to apply for that job.

A third method used in job recommender systems applies a *rule-based approach* to the rich information on jobs and workers on online job platforms to make recommendations based on the match between jobs and workers. The rule-based approach frames job recommendation as a classification problem and relies on worker's characteristics and job's requirements to predict the fit between the target worker and a certain job. For instance, if a worker satisfies the education requirement of a job, the job website is more likely to recommend this job to the worker.

Finally, some job boards apply more sophisticated systems that incorporate the hiring agents' behaviors into recommender systems and suggest jobs the target worker is likely to get feedback from (Kim, 2017). From the perspective of job boards, job seekers may become frustrated by sending out lots of applications but getting no echo, and switch to other sites as a result. Therefore, these *recruiter-behavior based algorithms* use recruiters' rating information to determine which type of jobs require which type of workers' characteristics and the probability of the worker getting callbacks when making job recommendations (Al-Otaibi and Ykhlef, 2012). More specifically, platforms collect the recruiter's application processing behaviors and predict the recruiter's preferences based on those behaviors. If the recruiter produces some positive signals towards a certain job applicant, such as browsing or downloading her resume, the system will acknowledge that the job prefers that type of job candidates, and recommend this job to workers who are similar to that job applicant (Yu et al., 2011). Moreover, the worker will receive job recommendations that are similar to this job since she has a relatively high chance to be suitable in similar positions.

## 3.4 Potential Mechanisms for Gender Bias

Although the algorithms used in job recommender systems are theoretically intended to be gender-neutral, there are at least four ways that recommender systems can deliver gender-biased job recommendations, which are connected to the four main components of most current job recommender systems.

The first is from item-based collaborative filtering recommendation, which recommends jobs that are similar to ones that the worker applied to in the past. While not in itself gender-biased, this algorithm tends to magnify and perpetuate previous gender differences in recommendations received by the worker. Suppose there is a job requesting male workers. In the extreme case, due to the gender mismatch, the job is not recommended to any female workers, and no female workers can see and apply for the job. In the following job recommendations, the absence of this job in female workers' application histories will reduce the exposure of other jobs that are similar to that job, even without gender request, and induce more divergence on the recommendation results between two genders.

The second component is content-based recommendations among workers. It is worth noting that the foundation of content-based recommendations, natural language processing algorithms can embody gender bias. For instance, female names are more associated with family than career words, compared with male names (Nosek et al., 2002). Nurse, teacher are more likely to be associated with she or her, while engineer, scientist are associated with he or him, suggesting that implicit gender-occupation biases are linked to gender gaps in occupational participation (Caliskan et al., 2017). If this is the case in job boards, we may observe some jobs are recommended to one gender more frequently than to the other gender because their characteristics are encoded to be correlated with gender identity, and the algorithm eliminates workers whose resumes do not contain the gender-related keywords (Savage and Bales, 2016). Furthermore, if the keywords associated with strong gender tendency are used to define similarity between workers, workers with the same gender consequently are more likely to be similar. For instance, patient is found in the resumes of female workers more often (or expatriate in male resumes), and if the algorithm uses these kinds of characteristics as the keywords in contents, workers are classified based on gender (Bozdag, 2013). As a result, female workers may be recommended with jobs that have been applied by other females, leading to gender segregation in job recommendations. Importantly, when jobs having gendered keywords are defined as similar, a worker that applies for one job with gendered words, will be recommended to other jobs that also contain such gendered words.

The third mechanism relates to the rule-based approach, which frames job recommendation as a filtering problem and only considers the 'hard' match between the worker and the job. If a job's characteristics are consistent with the worker's expectations and the worker satisfies the job's requirements, the job will be recommended to that worker. One important feature of Chinese job platforms is that they allow employers to explicitly state the gender of preferred applicants, without revealing these preferences to job seekers in the ads. Thus, for example, a rules-based algorithm might not show ads that list a preference for women to male job seekers.

Finally, consider recruiter-behavior based approach that incorporates hiring agents' rating behaviors (i.e., viewing and downloading profile, sending a message to target worker) into the recommender system. As far as I know, there are three scenarios in which the hiring agents' behaviors could affect job recommendations.[8] Suppose a hiring agent posted a job and received some applications from both genders, but

[8] Algorithms targeting at click maximization are likely to deliver biased results, due to the feedback loop (Jiang et al., 2019) and learning-to-rank approach (Jiang et al., 2016; de Sá et al., 2016).

has consistently ignored female applicants (for example, never downloaded female resumes).[9] Two points are learnt from this process: First, this job is not going to hire female workers, then it will not be recommended to other female workers. Second, if a female applicant did not get positive feedback from the job, the algorithm infers that she is unlikely to get callbacks from other jobs that are similar to that job, so those similar jobs will not be recommended to her. That is to say, workers' recommendation results are affected by the processing decisions of the hiring agents who posted jobs that they have already applied to, as well as the spillover effects from other hiring agents. Moreover, in most online job boards, hiring agents can search for and contact suitable workers directly. When a hiring agent searches for workers and clicks into a worker's resume, the jobs posted by this hiring agent will be recommended to that worker, as the hiring agent has shown interests to that worker (Köchling and Wehner, 2020). If a hiring agent persistently views resumes of male workers, those male workers will be suggested to apply while female workers do not have this priority (Burke et al., 2018).

The mechanisms mentioned above can interact with each other to create a complex job recommendation system.[10] More generally, algorithms may replicate the errors stemming from the training data, such as choosing parameters based on data with existing stereotypes, which detracts from gender fairness. Overall, recommender systems may reproduce and magnify pre-existing gender bias in the labor market.

[9] The four online job boards allow recruiters to filter workers' profiles by demographics (e.g., gender, age) and characteristics (e.g. education, experience) when they process received applications or search for suitable candidates.

[10] Both direct discrimination and indirect discrimination on gender potentially exist in these algorithms, which are distinguished by whether sensitive features (gender) are not explicitly used as inputs in algorithms (Pedreshi et al., 2008).

## 3.5 Experiment Design

### 3.5.1 Platform Environments

To cover a representative sample in online labor markets, the experiment was conducted on the top four job boards in China, which have millions of job seekers and job postings and can reach most of the workers and recruiters in the Chinese labor market. The large consumer bases allow me to create substantial fictitious workers but minimize the disturbance of the job search and recruiting process as well as the job recommender systems. The four job sites have similar interfaces and functions for users, with regular structures of online job platforms. Job seekers can register and create a profile for free, while employers are charged for posting job advertisements and using recruiter tools. Job seekers make applications by sending their resumes to the jobs that they are interested in, and hiring agents of firms can check and process the applications online and contact applicants through the website's message system. Furthermore, as far as I know, the leading job boards use more detailed and sophisticated forms of machine learning to suggest jobs to workers, and I may expect that the advanced algorithms may reinforce gender bias in an implicit way.

### 3.5.2 Job Type Selection

When a job seeker sets up her profile, job platforms let her indicate her current and desired industry and occupation. This job type information will be used by the job recommender systems and affect job recommendation results.

The selection of job types is based on three criteria: sample size, gender type, and hierarchy level. As a first step, I chose industry-occupation cells that have a large number of job postings to ensure that there were enough new job vacancies to be recommended to workers.[11] For instance, the internet industry has the most job postings, while sales are the most popular occupations in job sites, so the internet-sale is a potential job type. Second, because male-and female-dominated jobs might prefer applicants of different genders, I focused specifically on three gender types of jobs: female-dominated (i.e. administrative assistant), (approximately) gender-balanced (i.e. sales), and male-dominated jobs (i.e. software engineer).[12] Finally, because employers' gender preferences may also vary across the job ladder in which few women reach the top positions on the job ladder (Bertrand et al., 2010; Pekkarinen and Vartiainen, 2006), I diversify the hierarchy by including jobs in entry-level, middle-level and high-level. Taking the job of sales as an example, salesclerk is the entry-level job, sales manager is a middle-level job, and sales director is a high-level job. The details of these job types and the related characteristics of workers are described in Appendix C.1.1.

### 3.5.3 Resume Setup

I next created resumes that are qualified for the above jobs. The fictitious resumes come in pairs, and the two workers in each pair have identical backgrounds, except that one is female and the other is male. These resumes are quite sparse and contain only the mandatory information that is required to set up a worker profile to make sure that the recommendation results are not driven by other details. To achieve valid job recommendations, resume information is generated based on the real job ads and workers' resumes. For each job type, I scraped 50 job ads and 50 resumes as the information pool for fictitious profiles.

[11] The industry-occupation cell refers to the sub-industry and sub-occupation because workers will choose the finest category of industry and occupation when they set up their profiles.

[12] The selection of job's gender type is based on the public statistics and reports on the share of female workers in job boards.

A fictitious applicant's resume consists of four parts: personal information, education, job history, and job intention. Personal information section collects worker's name, birth date, years of working experience, current wage, city, employment status, phone number, and email address. Different from most audit studies relying on workers' names to signify gender identity, gender (male or female) is a compulsory input in Chinese job boards. The applicant's name is randomly assigned with the most popular names from 2015 Chinese Census 1% Population Sample, and the first name is matched with gender (See Appendix C.1.2 for more details). Since Chinese employers' gender preferences appear to interact strongly with the worker's age (Helleseter et al., 2020), I create two versions of each matched profile pair—a 'young' and an 'older' version, in which 'older' workers refer to ones who have more working experience. Worker's age, education and working experience are jointly determined. Young workers graduated in 2017, have three years of working experience, and are either 25 years old (born in 1995) if he has a college degree, which takes three years to achieve, or 26 years old (born in 1994) if he has a bachelor's degree, which takes four years to achieve. The corresponding older workers are 35 (college) or 36 (bachelor's) years old with 13 years of working experience. The specific education level and academic major satisfy the requirements of job type, and the school's name is randomly drawn from the Chinese High Education Institution List.[13] All the applicants are currently employed, and their wages are crafted to match the wages of existing job seekers by job type, education level, and years of working experience. As over half of job postings are from first-tier cities, I restrict the location of applicants to the first-tier cities in China, including Beijing, Shanghai, Shenzhen, and Guangzhou. Each applicant has a unique and active email address and mobile phone number.

In terms of job history, young workers started their current jobs in August 2017,

[13]Released by the Chinese Ministry of Education in 2019.

just after they graduated with the highest degree. For older workers, the beginning date of their current jobs is August 2015, implying that they have 5 years tenure in their recent positions. Worker's current occupation and industry are the same as the job type's occupation and industry, and job title and job description are entered as the job's occupation. I make up the company name to minimize the disturbance to both job seekers and employers in job platforms, which is a combination of worker's city, industry and a randomly generated name (i.e., Beijing Dongya Internet Technology Company). In the job intention section, a worker's desired wage is 120% of his current wage, and the desired city, industry and occupation is aligned with current ones.[14] Appendix C.1.2 summarizes the details of resume generation process.

To sum up, I created groups of four resumes that vary along two dimensions, gender and age, with all the other characteristics and information held constant, except that the older resumes' experience and current wages are adjusted to be age-appropriate. Given that the four workers in each group are designed to have the same job type and 35 job types are selected in each job board, I created 140 fictitious profiles (replicated across 4 cities) on each platform. After finishing the profiles creating process, male and female applicant published their profiles at the same time, afterwards their resumes are accessible (can be read or downloaded) to recruiters and headhunters on the platforms.

### 3.5.4 Implementation

In addition to workers' resume characteristics, job recommender systems user workers' browsing and application behaviors to deliver customized recommendations. To control for such differences, the paired (male and female) profiles followed identi-

[14] According to the salary reports from the job boards, 20% is normal and moderate wage growth for an average worker switching to a new job.

cal application strategies. Fictitious workers are naïve users on the job platforms, who click into and send resumes to the top listed recommended jobs. This process works as follows:

- **Round 0**. The male and female workers with newly created resumes log into their accounts at the same time, and I collect the first advertisement listed in the recommendation interface, to a maximum of 100 jobs. Then the workers log off.
- **Round 1**. The male and female workers simultaneously log into their accounts again, and I record the top 10 jobs (1st to 10th of listed job ads) in their recommendation lists. The two workers then apply to the top 10 recommendations by submitting their resumes. Immediately afterwards, the workers refresh their webpages and I record the 10 recommended jobs that appear.
- **Round 2**. At two-week intervals, I repeat the Round 1 procedures.
- **Round 3**. After two weeks, I repeat the Round 1 procedures again.
- **Round 4**. After two weeks, male and female workers log into account at the same time, and I record the number of views on the worker's resume by hiring agents. [15]

Figure 3.1 demonstrates the timeline of this experiment. Ideally, each fictitious worker received 160 recommended jobs and applied for 30 jobs in an 8-week job searching spell, and the collected outcomes include the information of 160 jobs as well as the number of hiring agents' views on profiles. The design of my field experiment guarantees that any observed differences in the job recommendations are caused solely by my randomized gender manipulation.

[15] After a worker's resume opens to the public, it can be viewed by all recruiters on the job boards. Recruiters of the applied jobs can read applicants' profiles, and other recruiters can find workers by searching resume, or by worker recommendations from job boards. The number of views records how many times that the resume is read by hiring agents.

## 3.6 Results

My audit study of job recommendation algorithms started in July 2020 and the last round of collections on hiring agents' views was completed in April 2021. In total, 2,240 fictitious profiles were created in four job sites, and those workers received 319,974 job recommendations from 119,356 individual job advertisements.[16]

Table 3.1 presents the descriptive statistics of my sample of fictitious workers. As applicants are designed in pairs and have fixed characteristics, Table 3.1 mainly reflects the presence of job boards in labor markets. The average annual wage of worker sample is 142,507 RMB, which is around twice the national average wage of workers in the urban in 2020.[17] The desired wage is 26.1% higher than the current wages,[18] and the average years of education is 15.56, indicating about half of the fictitious workers hold a bachelor's degree.[19]

The sample of recommended job ads is summarized in Table 3.2. Conditional on information is visible to workers, the average job posted an annual wage of 211,004 RMB, requested workers with 14.4 years of education and 2.3 years of working experience. On average, employers advertised a wage that was 17.4 percent higher than the fictitious workers' desired wages, but requested lower education levels and fewer years of

[16] There are several reasons that the recorded number of job recommendations is smaller than the designed number 2,240*160 = 358,400. The first reason is in Round 0, some job types did not have 100 active job openings that matched the characteristics of workers. The second reason is that, job boards froze suspicious workers' accounts and a few of them were blocked after Round 0. If one account in a gender pair was blocked, I terminated the experiment of the whole gender pair. Another reason is job ads were withdrawn by hiring agents when I scraped job ads so some jobs' information was unavailable. The missing data is less than 0.5% and occurs randomly, especially it is independent of the gender of fictitious applicants, thus unlikely to bias my analysis.

[17] According to the statistics from National Bureau of Statistics of China, the average annual wage of workers in the urban non-private sector in 2020 was 97,379 yuan (US$15,188), and workers in the urban private sector had an annual wage of 57,727 yuan (US$9,004).

[18] Some job boards let the worker choose desired wage range, and the desired wage is the midpoint of selected desired wage range.

[19] It is common to take 16 years to achieve a Bachelor's degree, and 15 years to achieve a college degree.

working experience. In addition, above 95% of job ads have explicit wage postings,[20] and one-third of the recommended positions are from companies that have more than 1,000 employees. The fictitious workers' profiles are well-matched with recommended jobs, as shown in Table 3.3. More than 80% of designed workers satisfied the jobs' requirements on education and working experience, and almost all of the recommended jobs' locations aligned with the worker's current location. 86.6% of recommended jobs posted wages that were higher than workers' lowest desired wages.

### 3.6.1 Set and List Differences between the Job Recommended to Men and Women

This section answers the most basic question about gender bias in job recommendations: To what extent are the jobs recommended to male and female workers the same, or different? I quantify the gender difference in job recommendations in two dimensions: the set difference and list difference.

#### The Set Difference

*Set difference* measures the share of jobs that are only recommended to one gender, without considering the sequence of recommended jobs. Figure 3.2a demonstrates the set difference: Suppose for all workers, male applicants receive jobs that are in set A and C, and the female applicants are recommended by jobs in set B and C, in which set C contains the overlapped jobs of female and male recommendations, while set A represents the only-to-male jobs, and set B includes the only-to-female jobs. Then the set difference rate is defined as the share of only-to-one gender jobs on the whole pool

[20]While some empirical evidence suggests that better jobs (i.e. higher requirements on education and experience) are less likely to explicitly post wages (Marinescu and Wolthoff, 2020), it is not true in my data.

of recommended jobs received by male and female applicants:

$$Set\ Difference\ Rate = \frac{\#\ jobs\ in\ A + \#\ jobs\ in\ B}{\#\ jobs\ in\ A + B + C} \tag{3.1}$$

I present the set difference rate by worker's age level, by job's gender type, by job's skill level and by city in Table 3.4. In total, the set difference rate between male and female applicants is 9.72%, meaning that out of 100 jobs recommended to male and female applicants, 90.28 jobs are displayed to all applicants, and 9.72 jobs are unique to one gender while applicants with the opposite gender cannot see those jobs in their recommendation lists.

While we expected that the gender difference in job recommendations would be greater in jobs typically occupied by males or females, our empirical results do not support that claim. In contrast, male and female applicants working in gender-neutral jobs observe about 1 additional different job per 100 recommended jobs, compared to workers in male- or female-dominated job types. For the age variation, young applicants who have 3 years of working experience are more likely to be exposed to gender-specific job ads, but the difference is quite small. Job hierarchy also matters job recommendations to men and women, in which gender-specific jobs appear more frequently in middle-level jobs, and least in entry-level jobs. The last panel of Table 3.4 shows the geographical evidence on the share of gender-specific jobs. Generally, the set difference rates are close across the four cities, indicating no spatial disparity on the gender difference in job recommendations is detected.

To confront the issue that the pattern of gender bias may vary across subgroups, I decompose the number of different jobs between two genders by age, job's gender type and hierarchy in Figure 3.3. Two features can be identified: First, Figure 3.3a illustrates that the gender difference in recommended jobs is greater for young pairs and more pronounced in gender-neutral jobs in older pairs. Second, in Figure 3.3b, female-dominated with middle- and high-level jobs contribute a large share of different jobs between male and female applicants. Figure 3.4 further displays the dynamics of set difference rate. Without making any job applications, the share of gender-different jobs in Round 0 is 7.66 percent, and the share goes up after Round 1 when workers started to make job applications and is perpetuated with the application process. At Round 3, the chance of applicants viewing a gender-specific job is more than doubled relative to the share in Round 0.

### The List Difference

While the set difference reveals the number of recommended jobs that are unique to one gender, it only partially uncovers the difference on job recommendations, because job ads are ranked in the recommendation list; ones displayed at the top receive more attention, and are more likely to be seen and clicked into by workers (Craswell et al., 2008; Richardson et al., 2007). For instance, if the jobs received by male and female workers are completely the same, but male workers observe the jobs ranked from high to low (in quality) and female workers see a list of an opposite order, we can hardly say that they achieve the "same" job recommendations.

Now I take the rank of recommended jobs into account to measure the gender inequality in job recommendations. Define two job recommendation lists are the same only if the two jobs in the same rank are identical, as shown in Figure 3.2b. Then the list difference rate is defined as:

$$List\ Difference\ Rate = \frac{\sum_{i=1}^{n} ith\ job\ ad\ is\ different\ in\ gender\ pair}{Length\ of\ recommendation\ list\ (n)} \tag{3.2}$$

Table C.7 summarizes the average list difference rates by experimental rounds, worker's age level, job's gender type, hierarchy level and city. The difference rate inflates after considering the ranks of jobs. The overall list difference rate is 70.7%, indicating that in a list of 100 recommended jobs, there are about 30 jobs that are displayed identically to male and female applicants. Similar to Figure 3.4, the list difference rate largely increases after applicants send out job applications, from 58.3% in Round 0 to 86.4% in Round 3. The list difference rate has a quite consistent pattern with set difference rate across the subsamples by age, job's gender type, hierarchy level and city.

### 3.6.2 Differences in the Quality of Jobs Recommended to Men and Women

As shown above, job recommendations to male and female workers are not the same. This dissimilarity does not necessarily indicate bias, however, because it could result from randomness in each website's recommender system. However, if systematic gender bias actually exists in job recommendations, jobs recommended to one gender would be better than jobs shown to the other gender. To address this question, this section explores whether job recommendations to the two genders are equally good.

#### Explicit Measures: Wage, Education and Experience Requirements

While jobs can be evaluated from various dimensions (Brenčič, 2012), I start from the explicit characteristics: the job's posted wage, requested years of education, and requested years of working experience. In order to compare the quality of different jobs recommended to male and female applicants, the subsequent analysis sample is composed of jobs unique to male applicants (i.e. jobs in Set A in Figure 3.2a) and jobs unique to female applicants (jobs in Set B), and the overlapped jobs (i.e. jobs in set C) are excluded.

I use the two-sample t-test to examine whether the mean of the characteristic in male-only jobs equals the mean in female-only jobs. Suppose in job ads received by all men and women, the observed job's characteristic $x$ in male-only job sample is $(x_1^M, ..., x_n^M)$, and in female-only job sample is $(x_1^F, ..., x_n^F)$, where $n$ denotes the number of different jobs in male or female job recommendations,[21] then the null hypothesis of two-sample t-test is:

$$H_0 : \overline{x^M} = \overline{x^F}$$

Taking the posted wage as an example, under the null hypothesis, the average posted wage of only-to-male jobs does not differ from the average posted wage of only-to-female jobs.

Table 3.5 presents the results of two-sample t-test on the job's posted wage, education requirement and working experience requirement.[22] Conditional on the wage is advertised publicly, the gender gap of recommended wage between male and female applicants is 2,709 RMB and significant at 10% level, which is equivalent to 1.9% of the average current wage of fictitious workers, meaning that jobs recommended to men propose higher wage on average the jobs recommended to women.[23] The requested education is statistically indistinguishable between male-only and female-only jobs, but the required working experience in male-only jobs is significantly higher than the requirement in women-only jobs by 0.08 years, which is translated into 0.5% of the average worker's working experience.

To facilitate comparison on the gender gap for subgroups, I provide the two-sample t-test results by experimental rounds, worker's age, job's gender type and hierarchy

[21] n can be different for male and female applicants due to the replicated recommendations.

[22] Equal variance is applied, and results for two-sample t-test with unequal variance are in Appendix Table C.6.

[23] Instead of an exact wage, most jobs posted a wage range. The job's wage in the analysis is the midpoint of posted wage range.

level in Figure 3.5. According to Figure 3.5a, the jobs' wages in male-only and female-only recommendations do not differ among young workers and older workers, across female-dominated, gender-neutral and male-dominated jobs, and across job levels, suggesting that gender effect on recommended wage does not interact with age and job type (gender and level). But regarding the rounds, in rounds 3, after workers applied 20 jobs, only-to-male jobs post significantly higher wages than only-to-female jobs. Similar to the previous results, the differences in education requests of recommended jobs to males and females are positive but remain insignificant in all subsamples, as shown in Figure 3.5b. In Figure 3.5c, the higher requirement on working experience in male-only jobs is pronounced in older applicants, in one gender-dominated jobs, and in entry- and high-level jobs.

Two-sample t-test assumes the variables are continuous and normally distributed or large sample size, and those assumptions might be violated in the analysis sample (i.e., the requested experience is ab integer). I provide the Wilcoxon rank-sum test as a robustness check in Appendix Table C.6, which is a non-parametric test without assuming the certain distribution of variables, and the main results do not alter under Wilcoxon rank-sum test.

Turning to the list difference, I construct the comparison between male and female job recommendations by using paired t-test. After excluding the identical recommendations (i.e., 1st and 2nd job recommendations in Figure 3.2b), a paired applicants' job recommendation lists can be expressed as $\big((Y_1^M, Y_1^F), ..., (Y_s^M, Y_s^F)\big)$, in which $s$ denotes the different recommendations. Suppose the $ith$ job recommendation is $(Y_i^M, Y_i^F)$, and $Y_i^M$ and $Y_i^F$ represents the $ith$ job recommended to male and female applicants in the list- different recommendation sample. Define the difference $d_i$ of job's characteristic $y$ in $ith$ recommendation as:

$$d_i = y_i^M - y_i^F \tag{3.3}$$

Under the paired t-test, the equally good job recommendations mean that the average difference between the characteristics in the two jobs listed in the same position is not significant from zero. The null hypothesis is:

$$H_0 : d_i = 0$$

Compared to two-sample t-test, paired t-test assumes that the two jobs recommended to male and female applicants in the same rank are correlated. Table C.8 replicates the computation in Table 3.5 but replaces two-sample t-test with paired sample t-test. On average, jobs recommended to male and female workers in the same rank do not differ in their posted wages, required education and working experience.[24]

**Implicit Measures: Words**

In addition to explicit measures, a job's quality can be measured using the words in the job descriptions. Moreover, the presence or absence of a specific word may affect the matching between workers and jobs and leads to gender segregation in job recommendations (Dreisbach et al., 2019). In this section, I explore the gender difference in the wording in recommended jobs.

In a typical job ad, the job description is one or two paragraphs of text, which is placed after the explicit characteristics of jobs and contains rich information about the position. While the contents of job descriptions in different job types are highly diverse, they can be broadly aggregated into five categories:

(1) *Skills.* Skills are the core part of the job description, and recruiters express skills in various ways. For example, skill is often stated as a job requirement, "the candidate

[24]Similar to two-sample t-test, paired t-test requires that the measured differences are continuous and normally distributed. A robustness check with Wilcoxon signed rank test is presented in Appendix C.3.

should be familiar with Excel", or a part of the position description, "common tasks include making reports with Excel". While plenty of methods are developed to deal with the complexity of skills in jobs, I adopt the skill classification in OECD Programme for the International Assessment of Adult Competencies (PIAAC) (OECD, 2016), which is widely used in the social sciences research on gender differentials on skills (Christl and Köppl-Turyna, 2020; Pető and Reizer, 2021). More specifically, skills are divided into seven subsets, including literacy skills, numeracy skills, information and communication technology (ICT), problem-solving skills, influencing skills, co-operative skills and self-organising skills.

(2) *Benefits.* In addition to offered wage, employers advertise jobs' benefits to attract applicants. In Chinese job boards, the advertised benefits are often tagged, and their expressions are quite uniform across job types and platforms. Based on the extracted information from job ads, I classify job benefits into four types: payment, break, facility and insurance.

(3) *Work form.* Work form is a wide category that introduces the working time arrangement, capturing words about work schedule, business travel, work break and work overtime.

(4) *Company.* Job ads provide information on both the position and the company. The company features are summarized into three parts: workplace environment, company type, and title.

(5) *Other requirements.* Instead of simply displaying education level and years of working experience, employers state more detailed requirements in the job description, for example, the prospective candidates study in certain academic majors, have overseas working experience or have certain personalities. The other requirements category refers to the aspects of desired worker's age, appearance, personality, education, working experience, and other conditions.

Based on the above structure, the information in job descriptions was extracted in the following way: For all the jobs collected from four job boards, I first segmented a chunk of text into words (phrases) and retained words (phrases) with high frequency. Then I combined the words (phrases) that have the same or close meaning (i.e., leadership vs leading) to make the selected words (phrases) clearly contrast with each other, and assigned them to one of the five categories. In total, I extracted 167 individual words from job ads, listed in Appendix Table C.9.

Figure 3.6 presents the word cloud of job descriptions, with the bigger size representing a higher frequency of words in job ads.[25] Words related to job benefits, such as insurance, vacation and payment scheme, are most commonly seen in job descriptions, and employers often state their requests on worker's communication skills, coordination skills, teamwork skills and leadership.

When the job recommendation is gender-neutral, the proportions of containing a certain word or phrase in male-only and female-only jobs should be close to each other, whereas when gender bias exists, some words will be differentially present in advertisements for jobs recommended to men and women. This hypothesis is examined by the proportion test: Word $z$ is constructed as a binary variable, and $z_i$ takes the value of 1 if job $i$ contains the word $z$ in its description. Under the proportion test, the null hypothesis is that the probability of the word showing up in male-only jobs equals the probability that it appears in female-only jobs.

$$H_0 : \overline{z^M} = \overline{z^F}$$

Table 3.6 displays the proportion test for wording difference in job recommendations under meaningful categories. The coefficient in parentheses represents the gen-

[25] Appendix Figure C.1 shows the word cloud in Chinese.

der gap in words frequency ($\overline{z}^M - \overline{z}^F$), and a positive difference means that jobs recommended to male workers are more likely to contain that word in their job descriptions than female workers' jobs. The left panel lists 28 female words, which have a higher probability of being included in female-only jobs at 5% significance level, and the right panel includes 31 male words that are significantly mentioned more in male-only jobs.

We can see that most literacy skills, such as *speak* and *documentation* are more common in only-to-female jobs. Furthermore, female applicants are recommended for more jobs mentioning *data*, *chat tools* and *administrative tasks*, while male applicants see more jobs that require *problem-solving skills*, such as *decision-making*, *engineering*, and *working independently*, and *influencing skills* such as *leadership* and *manage*. These findings coincide with the results from previous literature on the gender gap in skills that document women tend to carry out more executed tasks, less skill-intensive tasks and use their cognitive skills less than men (Pető and Reizer, 2021; Black and Spitz-Oener, 2010).

In the *work form* panel, jobs with *regular working hours* or *flexible* schedules are more likely to be recommended to women, and jobs with decreased flexibility, such as *overtime working*, *night work* and *commute*, are more likely to be recommended to men. This is in line with the finding that women are more willing to pay for flexible work arrangements (Flory et al., 2015; He et al., 2021; Mas and Pallais, 2017; Bustelo et al., 2020). For *benefits*, female-only jobs are more likely to mention *base pay*, *marriage leave*, *maternity leave*, *social security*, *unemployment insurance* and *parental leave* in their descriptions while only-to-male jobs emphasize providing *shuttle*, *medical insurance*, *vacation*, *free meal*, *reward* and *stock*. Orientation *training* is mentioned more in female-only jobs, while jobs from the *publicly-listed* companies are more often to be recommended to male applicants. In addition to skills requirements, jobs recommended to men request workers who are *self-motivated*, *innovative*, *experienced*, and able to handle work *pressure*.

Words associated with physical appearance, such as *figure*, *facial*, and *temperament*, and words about the feminine personalities such as *careful*, *patient*, *punctual*, *outgoing* and *trustworthy*, are more frequently emerge within female-only job advertisements. Jobs recommended to women also open to hire *new graduates* and workers *without working experience*, and request that the qualified applicants should be *healthy* and *below 35* years old. [26]

### 3.6.3 Words and Gender Stereotypes

The different presence of words in gender-specific jobs established above provides initial evidence that gender bias in wording exists in job recommendations. Previous research has shown that the wording in job advertisements reveals employers' preference on gender and would direct workers' application behaviors, even when employers do not make explicit gender requests. For instance, women found jobs less appealing when the job advertisements included more masculine wording (Gaucher et al., 2011), and feminine wording in job titles and job descriptions increases the share of female applicants (Kuhn et al., 2020; Chaturvedi et al., 2021). Job platforms may mediate employers' gender preferences in a special way in which the recommender systems target the desired workers by linking gender to the words in job descriptions, which can reinforce gender stereotypes and result in gender-based job recommendations. In this section, I further explore the relationship between gender stereotypes and words used in gender-specific jobs.

I rely on multiple external sources of data to identify the femaleness and maleness of words in job ads. The first source is the previous literature on gendered words,

[26] The list difference in word frequency between male and female job recommendations was explored through McNemar's Test, and differences of the word usage between jobs recommend to only males and only females are insignificant in most cases.

which refer to masculine and feminine words that are associated with gender stereotypes. While linguists focus on the commonly used words in daily life and the effect of gendered words on people's behaviors (Fitzpatrick et al., 1995; Gastil, 1990; Lindqvist et al., 2019), researchers in political science (Roberts and Utych, 2020) and in psychology (Bem, 1981; Hoffman and Hurst, 1990; Rudman and Kilianski, 2000) specify gendered words in various application scenarios and argue that the usage of gendered words would shape people's attitude and support to social values. Most of the gendered words identified by these literatures are adjectives, which describe men's and women's personalities (e.g., masculine words: confident, aggressive, strong vs feminine words: sensitive, kind, beautiful). More relevant to the current context, three papers have encoded the gendered words used in job advertisements and demonstrated the subsequent labor market outcomes. Gaucher et al. (2011) collected masculine and feminine words from published lists of agentic and communal words, and masculine and feminine trait words. Given the existence of jobs with explicit gender requests in developing countries, Kuhn et al. (2020) and Chaturvedi et al. (2021) used text analysis and machine learning techniques to predict the implicit maleness and femaleness for individual words in job ads, which provides gendered words about both worker's personalities and required skills. More specifically, Kuhn et al. (2020) apply the naïve Bayesian classifier to identify the likelihood of an explicit gender request based on the words in job titles in a Chinese job board, and Chaturvedi et al. (2021) make use of the text contained in detailed job descriptions in India and construct measures on whether the job ad text is predictive of an employer's explicit male or female preference using a multinomial logistic regression classifier. In this paper, I combine the gendered words from the above studies to create a list of male and female words in my job ads.

As an alternative approach, I carried out two surveys to collect people's perceptions about stereotypically male and female words in job ads. In the English version of the survey, I recruited participants from Amazon Mechanical Turk (MTurk), and let them rate words on maleness and femaleness. A corresponding Chinese version was conducted on Chinese workers. The question takes the same form for each word: "Suppose you are a recruiter and craft a job advertisement containing the following word, you tend to hire (a) no gender requirement, (b) men, (c) women". In this setting, people would perceive the ideal gender of the candidate for jobs that are masculinely or femininely worded. Details on the surveys are provided in Appendix C.4.1 and C.4.2.

The heatmap in Table 3.7 demonstrates the results on the consistency of stereotypical gender roles in words from these three approaches: previous literature, Mturk survey, and Chinese survey. The displayed male and female words are the ones that were identified by my audit study (from Table 3.6), and the color intensity represents the femaleness or maleness defined by three approaches. If a word is highlighted with bright red, it is defined as a female word in all three approaches. Words in light red are defined as female words in two approaches, and pink color means a female word in one approach. Male words are marked with blue colors, in which bright blue, light blue and pale blue represent maleness from three, two and one approach, respectively.

Overall, red colors in the left panel and blue colors in the right panel clearly demonstrate that male and female words, which emerge with significantly different frequency in male-only and female-only jobs, are correlated with gender stereotypes. Jobs that are only seen by men contain greater words describing male characteristics, such as engineering, leadership, and overtime, while women are more likely to be exposed to the ads including assist, administrative, patient and temperament, which are also conformed to female stereotypes. The wording in the job descriptions may convey information on the job's implicit gender requests through job recommender systems that encode gender with words reflecting workplace stereotypes on men and women.[27]

[27]This is consistent with results from Chaturvedi et al. (2021), who found that words related to hard-

Since gendered words occur differentially in gender-specific jobs, I further ask that, among these words, which of them can predict whether the job is a male-only job or a female-only job? I use four methods to measure the relation between words used in job descriptions, and whether it is an only-to-male job.

The first and basic method is OLS regression, in which the outcome variable is 1 if it is a job only recommended to male applicants, and 0 if it is a female-only job. The regressors are dummy variables for the presence of 167 words. Column 1 in Table 3.8 lists the top 10 words in magnitude that are significant at 5% level. and the overall F-test result is F(167, 24921) =2.63 ($p<0.0001$), indicating that the 167 words are jointly significant. As the matrix is large, sparse, and some of the words are correlated with each other, one may want to select variables that have a larger impact on the outcome rather than including all of them. I use lasso and ridge regressions that impose a penalty parameter for adding an extra variable to figure out which words correspond to the different recommendations to men and women. I applied 20-fold cross-validation to find the optimal penalty parameters, and the selected top 10 words by lasso and ridge regressions are shown in column 2 and 3.

The last approach to identify words that contribute to the classification of jobs recommended to men and women is a machine learning method, random forest. Given the binary measures for the outcome and independent variables, my data structure is very suitable for adopting a decision tree method to find the important factors that affect the sample splitting to male-only and female-only jobs. Column 4 in Table 3.8 presents the top 10 words with high feature importance based on 100 decision trees and Gini impurity. I find that the three regressions' results are quite consistent. For instance, marriage leave, base pay, words about working hours and breaks are highly predictive of gender-specific job recommendations. Random forest results suggest that

skills and flexibility are critical in explaining gender disparities in labor market outcomes.

words related to breaks, holiday and vacation, are important in making a job ad more or less male recommended.

Finally, to achieve an overall evaluation of gender bias in words, I compute the vector dissimilarity between the average jobs recommended to men and women. Based on the extracted words in the job descriptions, job $i$ can be expressed by a vector $S_i$ with 167 elements, in which the $ith$ word $s_{ij}$ $(j = 1, ..., 167)$ equals 1 if job $i$ contains word $j$. The dissimilarity between the average male-only job, $\bar{S}^M$ and the average female-only job $\bar{S}^F$ is computed as Euclidean distance between two vectors, and is plotted in Figure 3.7. It suggests that on the aggregate level, wording in jobs to men and women has a dissimilarity about 0.3, and gender-specific jobs recommended to young workers in male-dominated jobs and entry-level jobs have a slightly higher dissimilarity than such seen by older workers in gender-neutral jobs and middle-level jobs.

## 3.7 Explanations for the Gender Bias in Job Recommendations

Section 4 pointed out four mechanisms that could deliver gender-biased job recommendations. In this section, I attempt to distinguish between these reasons, in order to isolate which ones account for that bias.

First, my findings suggest that item-based CF enlarges gender bias in the application process, at least in part. Because the job recommender systems absorb workers' rating behaviors and suggest jobs that are similar to the previously applied jobs, recommended jobs would be more diverse when workers have different application histories. I isolate the impact of item-based CF on gender difference by comparing how the recommended jobs change before and after making applications. Quantitatively, according to Figure 3.3, the set difference rate rises remarkably after applicants send out profiles in Round 1. Figure 3.8a illustrates the gender gaps of explicit measures on jobs' quality increase after applications, which is particularly true for requested working experience. For the wording in job ads, Figure 3.8b shows that on the aggregate level, word dissimilarity between male- and female-only jobs increases after workers start to apply. Furthermore, by checking the number of male and female words (defined by proportion test with 5% significance level), I find that gendered words in gender-specific jobs increase after applications: The number of male words increases largely from 8 to 19, and the number of female words increases from 18 to 23.

Secondly, the association between gender and words established by content-based recommendations plays a role in gender-biased job recommendations. On the basis of findings from Table 3.7, female words and male words contained in gender-specific jobs are correlated with gender stereotypes in the workplace. For instance, *figure*, *patient*, and *maternity leave* are feminine-themed words, which also have a higher frequency in the female-only jobs, while jobs recommended to males involve more maleness words such as *engineering* and *leadership*, implying that gender-related words may be encoded and applied into the job recommendations.

Thirdly, rule-based approach that complies with employers' stated gender requests probably has a very limited effect on gender-biased recommendations. While I cannot observe the preferred gender from public job ad postings, recent studies show that jobs advertised specifically for men or women have substantially reduced due to the recent policy interventions from the Chinese government (Kuhn and Shen, 2021). In Kuhn and Shen (2013), jobs that specified desired gender accounted for about 10.5% in Zhaopin in 2008, while my internal data from Liepin suggests that the share was lower than 1% in 2018. Moreover, if the gender requests still exist, they are more likely to appear in the fields that are dominated by one gender, thus we expect to find greater gender bias in male- and female-dominated jobs. However, my decomposition on the setbeahvior difference rate as well as the explicit measures for the quality of gender-specific jobs implies that there is no strong evidence that applicants in male- or female-dominated jobs received more gender-specific job recommendations, or the gender disparities in job's quality magnified in those fields.

Finally, other features of my results suggest that recruiter-behavior based algorithms affect job recommendations. Although my fictitious profiles are very brief and rarely get callbacks from employers, the number of profile views are recorded by websites, including the views from hiring agents who process the received applications, as well as the views from other hiring agents who find the worker's resume through search function or worker recommendations in the job board. If a hiring agent shows interest in a certain worker, then the worker may be recommended to apply for jobs posted by that hiring agent, indicating that the human bias might be manifested and reinforced by the algorithm bias.

To prove this claim, I run the regression of the set difference rate on the number of views on the male and female profiles on gender pair level:

$$Y_i = \beta_0 + \beta_1 ViewF_i + \beta_2 ViewD_i + AX + e_i \tag{3.4}$$

The outcome variable $Y_i$ is the number of different recommended jobs per 100 recommendations in gender pair $i$ (100*set difference rate), and the variables of interest are the number of views on female profile, $ViewF$, and the gap of received views between female and male profiles in the gender pair, $ViewD$. Table 3.9 reports the regression results. Column 1 only includes the two measures of views on gender pairs. Column 2 and 3 add controls for worker's age and job's gender type. In column 4, I further control for job board fixed effects to absorb various behaviors of hiring agents in different job boards. The estimation results show that views of the female profile are a significant contributor to the quantity of different jobs seen by identical men and women. The effect of the gender gap in views on the share of gender-specific jobs remains insignificant, however.

It is worth noting that although item-based CF, content-based recommendations, and recruiter-behavior based algorithms potentially generate and perpetuate the gender bias in job recommender systems, I cannot rule out the interactions between those channels, for instance, and several mechanisms can lead simultaneously to the biased results.

## 3.8 Conclusion and Discussion

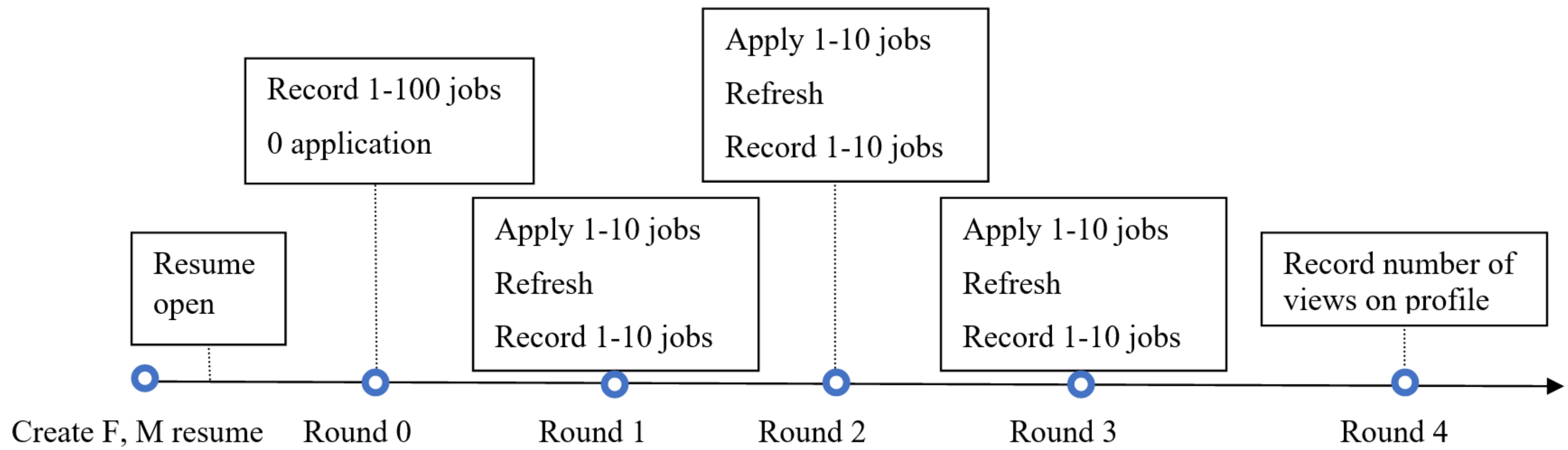
Computer scientists have proposed a variety of ways to improve fairness in algorithms. Most of these approaches focus on enhancing the algorithms' design using computations and formulas to minimize the risk of unfair treatment of certain groups of people (Bozdag, 2013). Recently, a strand of economic studies has attempted to eliminate algorithmic bias by introducing economic concepts into the algorithmic predictions. For instance, Kleinberg et al. (2018) point out that blinding algorithms to the candidate's identity is not a panacea for eliminating biases. They also argue that the inclusion of social planner who cares about equity in the prediction model can promote algorithmic fairness. In studies on particular applications, Arnold et al. (2021) propose approaches to measure discrimination in algorithmic predictions in the context of pretrial bail decisions, and Mullainathan and Obermeyer (2021) consider the label choice bias in algorithms. Moreover, Rambachan et al. (2020) focus on the prediction policy problems and address how to establish the optimal algorithmic regulation from the perspective of economics.

In the specific case of recommender systems, sophisticated recommendation algorithms have proven to be effective in supporting human decisions in discovering new items and are broadly applied in various fields. However, users, sometimes even the designers, have limited knowledge of the recommendation generation process, and the 'black box' of recommender systems may inadvertently cause problems in real social and economic life. However, empirical evidence is still in its infancy in this field. My paper provides an example through an algorithm audit to assess the causal effect of gender in job recommender systems. Using both set differences and list differences to measure the gender gap, I find that identical male and female applicants received different recommendations, in which women were more likely to see low-wage jobs requesting less working experience, requiring literacy and administrative skills, and containing words related to female stereotypes than comparable men. With the growing use of online job searching and recruiting, further research on gender differentials in labor markets should take the job recommendation bias into account.

Since the main objective of the recommendation system is to accurately predict users' (i.e., job seekers on internet job boards) interests, the objective of fairness is possibly overlooked and fails to be incorporated into such systems (Sonboli et al., 2021). While recently some researchers have proposed the fairness-aware recommender systems, it remains an open challenge given that fairness is difficult to define, track, and validate in recommendation systems in which every user expects a different item list based on her taste (Ge et al., 2021; Gao et al., 2021; Beutel et al., 2019; Fu et al., 2020). In my context, job recommender systems that are free of gender bias should theoretically ensure that male and female workers with the same qualifications get recommendations of jobs with the same quality. But what if men and women behave differently in job search such that men are more likely to click optimistically on high-paid jobs than women (Burke et al., 2018)? Furthermore, on *multi-sided* platforms, fairness and utilities of *all* stakeholders should be considered. When hiring agents' feedback influences the recommendation results, should the recommender system truly reflect employers' preference on desirable workers to facilitate potential job-worker matches even if the human decisions are biased? If fairness involves showing workers many jobs they have almost no chance of getting, is that desirable? In the long run, how to maintain the algorithmic fairness when new variables are introduced in the dynamics of job applications? What are the principles to make adjustments or corrections when the bias is detected in the system? All these questions remain unanswered and are good candidates for additional theoretical and empirical research.

Due to data limitations and the high complexity of job recommender systems, it is difficult to find the exact reason or sole driver for the gender bias in job recommendations from the observational data in algorithm audits (Hannák et al., 2017). More importantly, how the gender-biased job recommendations affect job seekers' searching outcomes is still masked. I hope that future research using field experiments or internal data from platforms will shed more light on those questions, and provide additional insights for anti-discrimination policy and legislation.

Figure 3.1: Timeline of the Experimental Steps




Notes:

1. Two profiles in each gender pair follow the same timeline. From Round 1 to Round 3, fictitious workers apply for the first job to the 10th job that are displayed in their customized job recommendation interfaces, and the time interval for each round is two weeks.

Figure 3.2: Difference Measures in Job Recommendations

Figure 3.2a: Set Difference

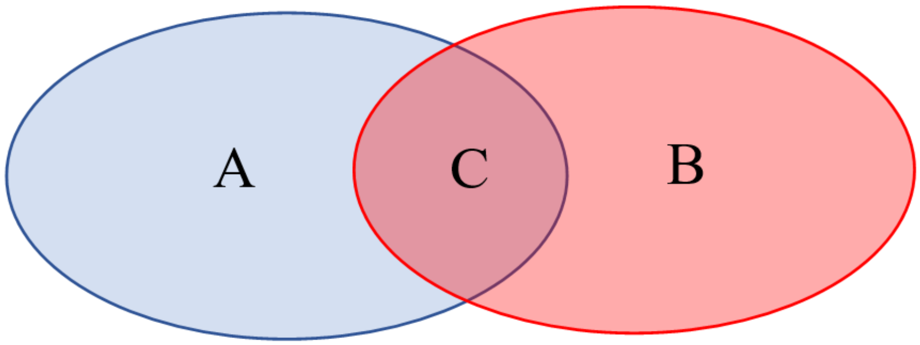



Figure 3.2b: List Difference

| | Male | Female | |
|---|---|---|---|
| 1st | Job 1 | Job 1 | Same |
| 2nd | Job 2 | Job 2 | Same |
| 3rd | Job 3 | Job 4 | |
| : | | | |
| ith | Job i | Job i+1 | |
| : | | | |
| nth | Job n | Job 3 | |

Notes:

1. In (a), Set A represents jobs that are only recommended to male applicants, set B represents jobs that are only recommended to female applicants, and set C represents the jobs that are recommended to both males and females. The set difference rate is defined as the share of gender-specific jobs on the complete set of recommended jobs, (A+B)/(A+B+C).

2. In (b), the shadow area indicates the identical recommendations in the gender pair, in which the $ith$ recommended job in the recommendation list of pairwise male and female applicant is the same. The list difference rate is defined as the share of different recommendations in the recommendation list. In the above example, only the first two jobs in recommendation lists are the same, then list difference rate is (n-2)/n.

Figure 3.3: Set Difference Rate by Age, Job Gender Type, and Job Hierarchy

Figure 3.3a: Set Difference Rate by Job Gender Type and Age

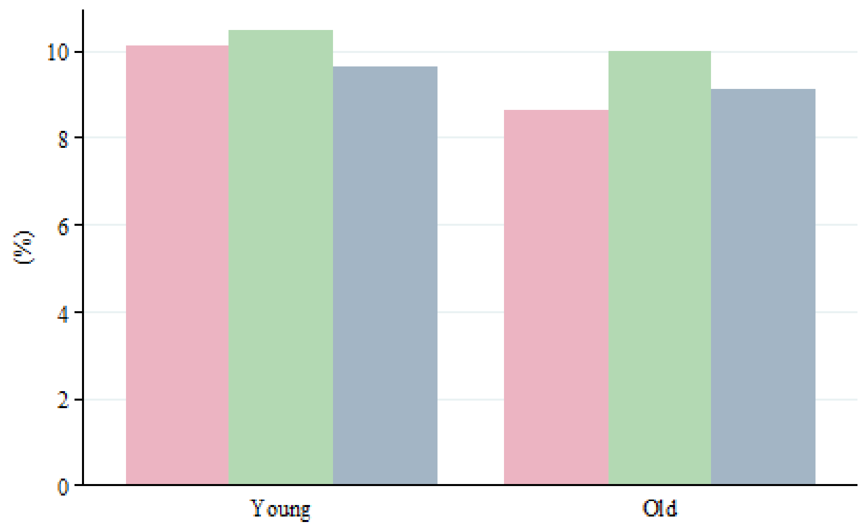
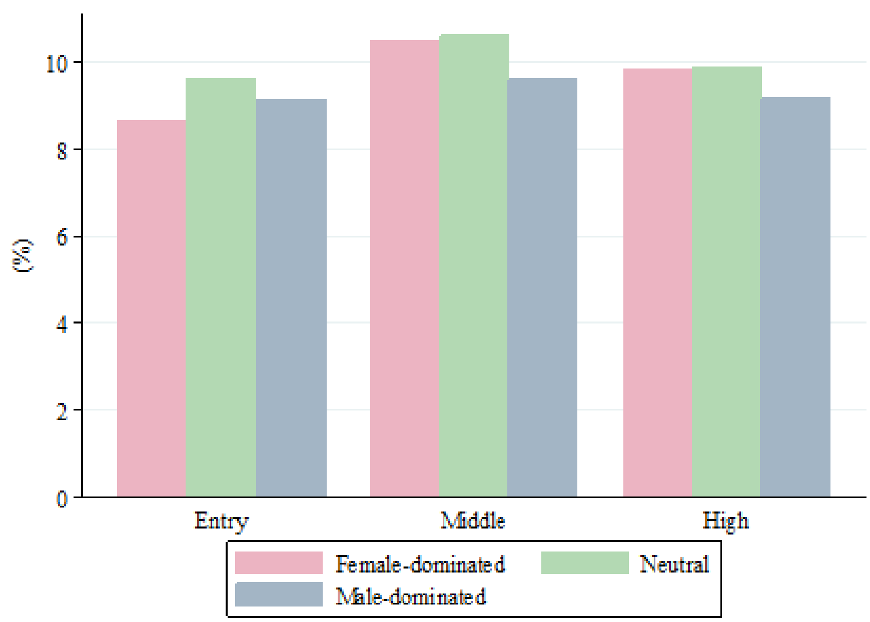


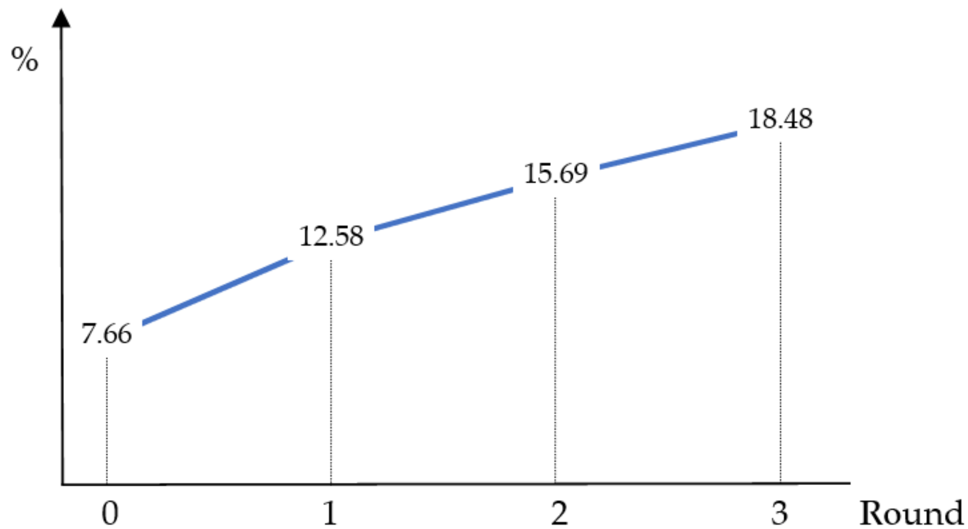

Figure 3.3b: Set Difference Rate by Job Gender Type and Hierarchy




Notes:

1. Set difference rate is defined on each group level. For instance, the first bar in (a) is the share of gender-specific jobs on the total jobs recommended to young workers in female-dominated fields.

Figure 3.4: Set Difference Rate by Experimental Rounds




Notes:

1. The number of job recommendations in Round 0 to Round 3 is 100, 20, 20, 20, respectively.

Figure 3.5: Gender Differences on Explicit Measures by Groups

Figure 3.5a: Gender Differences on Posted Wage by Groups

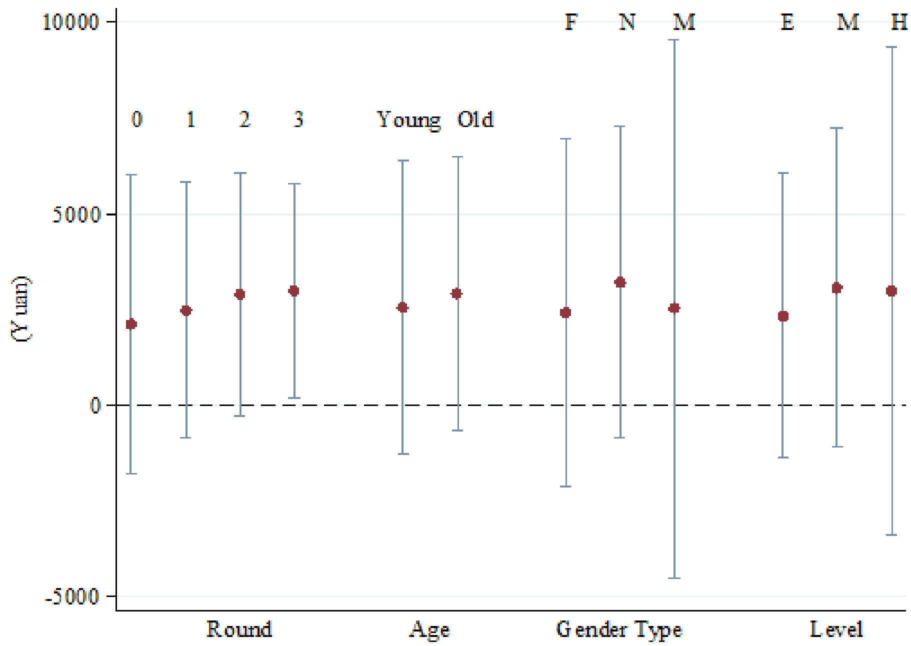


Figure 3.5b: Gender Differences on Requested Education by Groups

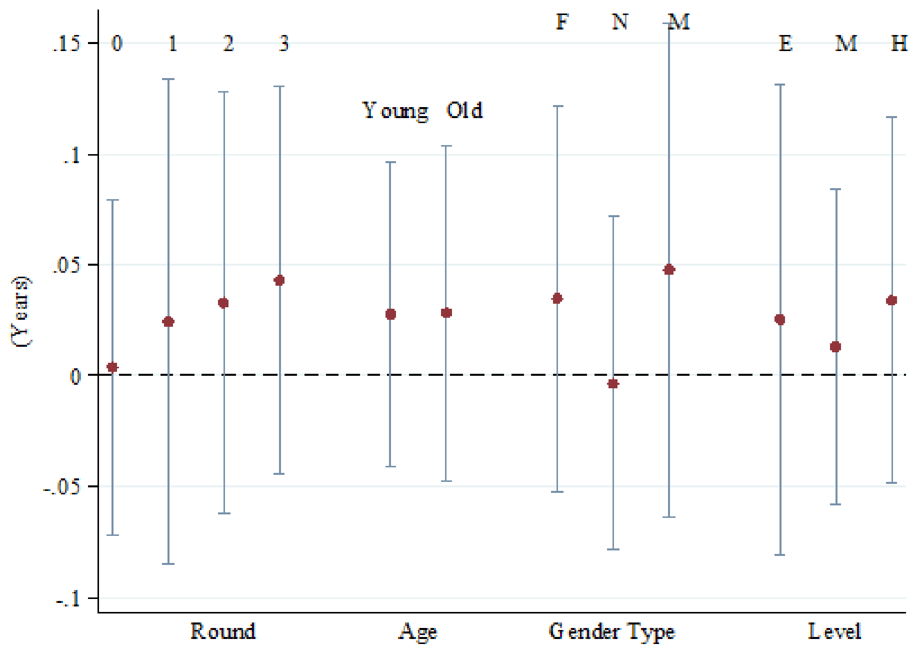
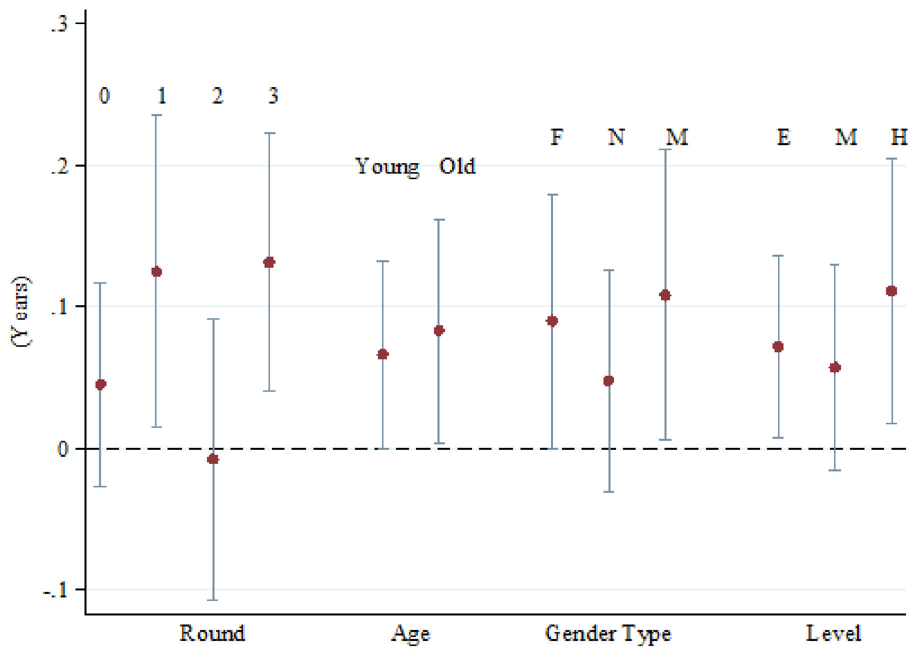


Figure 3.5c: Gender Differences on Requested Experience by Groups




Notes:

1. In the job gender type, F denotes female-dominated jobs, N denotes gender-neutral jobs, and M denotes male-dominated jobs. In the job hierarchy level, E denotes entry-level jobs, M denotes middle-level jobs, and H denotes high-level jobs.

Figure 3.6: Measure of Words' Dissimilarity in Job Recommendations


Notes:

1. The word cloud is based on the extracted words in the job descriptions from 119,356 recommended job advertisements, and the size corresponds the word frequency. The Chinese version is shown in Appendix Figure C.1.

Figure 3.7: Word Cloud from Job Ads


Notes:

1. The vector of the average only-to-male (female) jobs consists of 167 elements, in which each element represents the average frequency of that word in the only-to-male (female) sample. Dissimilarity is defined as the Euclidean distance between the male and female vectors.

Figure 3.8: Comparison of Gender-Specific Jobs Before and After Applications

Figure 3.8a: Comparisons on Explicit Measures

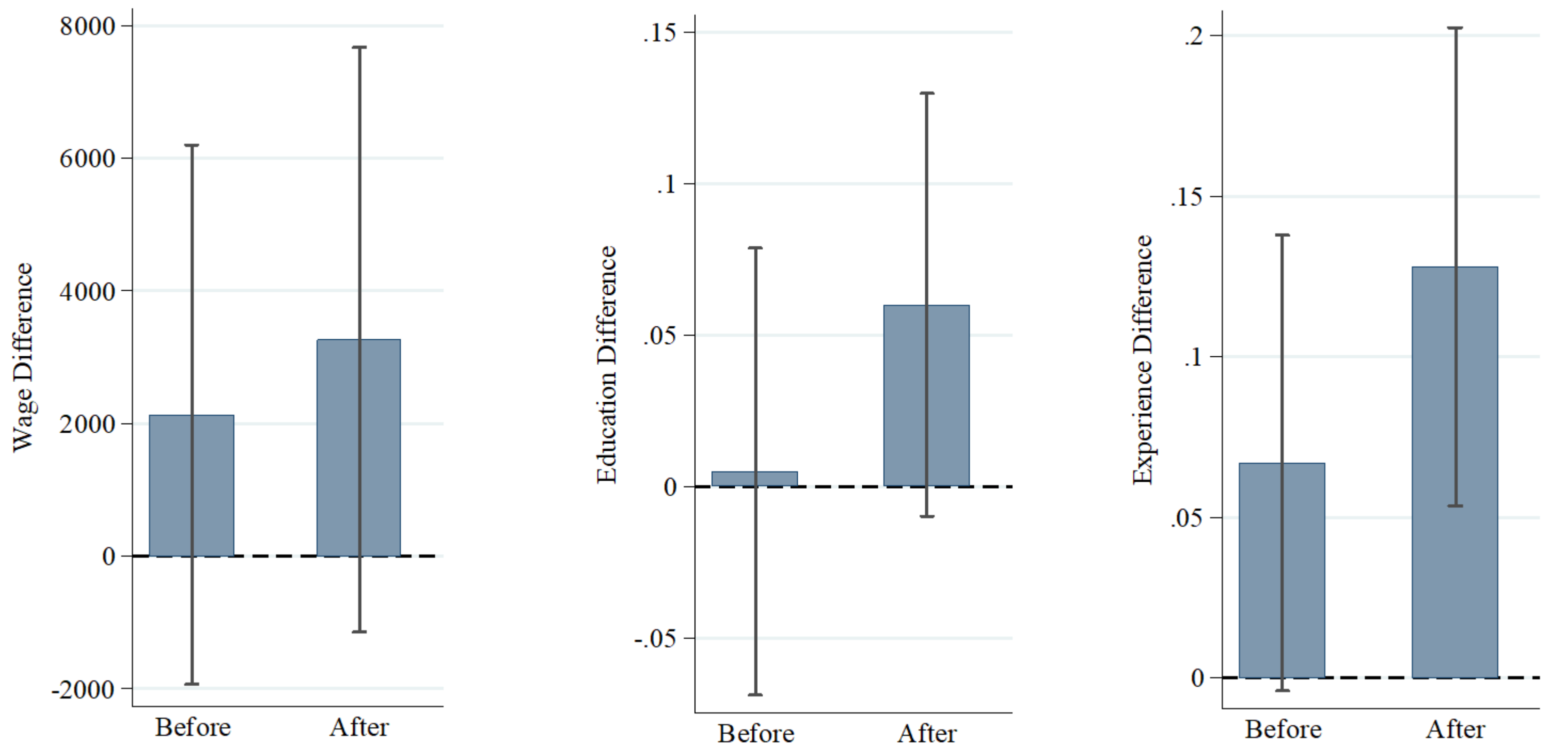
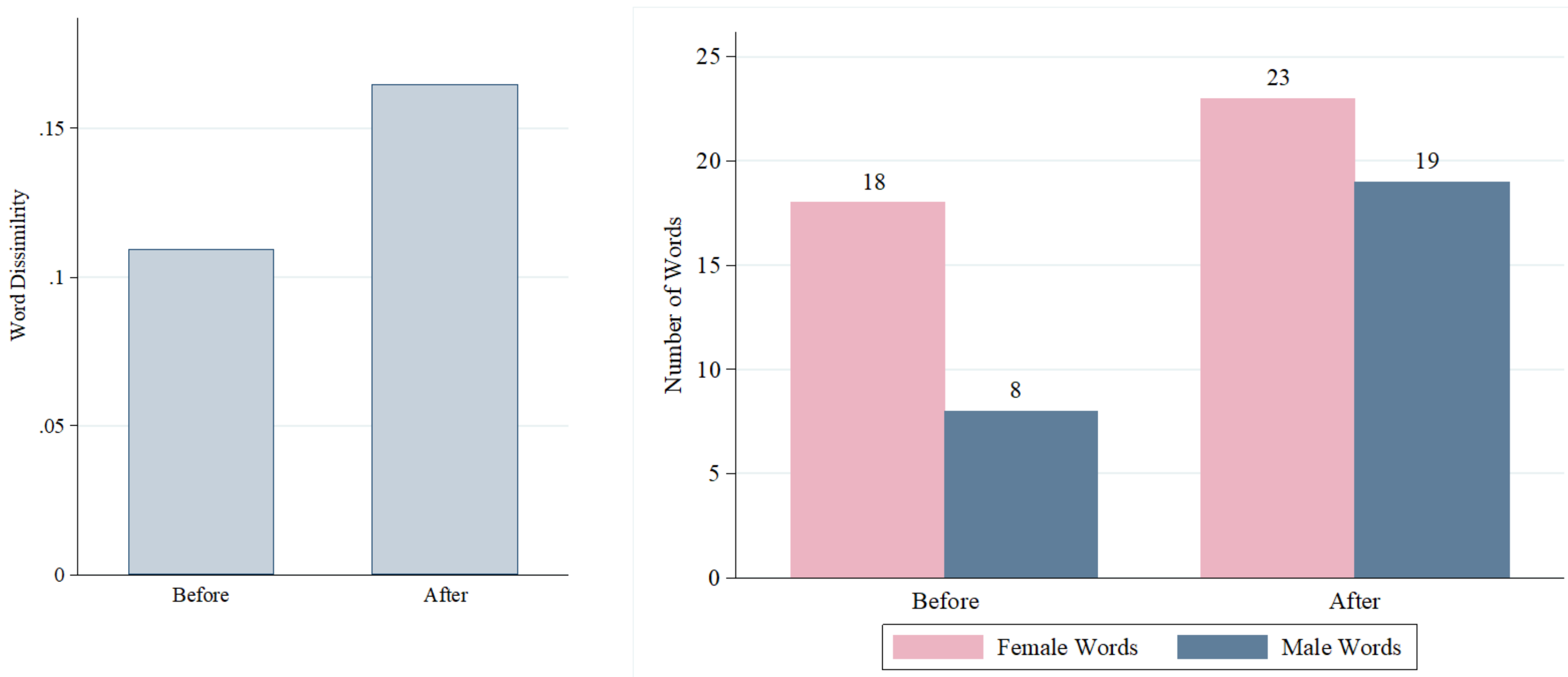


Figure 3.8b: Comparisons on Gender-Specific Words




Notes:

1. Figure 3.8a shows the gender gap of explicit measures on gender-specific jobs that are recommended before and after applications separately.
2. Figure 3.8b compares the wording in gender-specific jobs before and after applications. Word dissimilarity is defined as it in Figure 3.7, and the number of gendered words in job ads is computed with the same method in Table 3.6.

Table 3.1: Descriptive Statistics: Applicant Sample

| | Mean |
|---|---|
| Current Wage | 142507.1 |
| | (65141.8) |
| Desired Wage | 179732.1 |
| | (81818.9) |
| Education | 15.56 |
| | (0.4960) |
| | |
| Sample Size | 2,240 |

Notes:

1. Current wage and desired wage are annual wage in RMB.
2. Education levels in resumes are transformed to the years of education. A college degree is equivalent to 15 years of education, and a bachelor's degree is equivalent to 16 years of education.
3. Standard errors are in parentheses.

Table 3.2: Descriptive Statistics: Recommended job Sample

| | Mean |
|---|---|
| Posted Wage? | 0.9569 |
| | (0.2032) |
| Wage, if posted | 211004.3 |
| | (658266.8) |
| Required Education | 14.816 |
| | (2.1684) |
| Required Experience | 2.3082 |
| | (2.1119) |
| Large Company | 0.3564 |
| | (0.4789) |
| | |
| Sample Size | 119,536 |

Notes:

1. Wage is the midpoint of the posted range of wages.
2. Education levels in job ads are transformed to the years of education. Middle school takes 9 years of education, tech school and high school are equivalent with 12 years of education, college is 15 years of education, and bachelor's degree is equivalent with 16 years of education, master/MBA is 18 years of education, and doctoral degree is 23 years of education.
3. Large company refers to companies that have more than 1,000 employees. The company size is self-reported by hiring agents.
4. Standard errors are in parentheses.

Table 3.3: Descriptive Statistics: Job Recommendation Sample

| | Mean |
|---|---|
| Desired Wage Match | 0.8658 |
| Education Match | 0.8812 |
| Experience Match | 0.9183 |
| Location Match | 0.9924 |
| | |
| Sample Size | 319,974 |

Notes:

1. Desired wage match equals 1 if the recommended job's upper bound of posted wage range is higher than the worker's lowest desired wage.
2. Education (experience) match is 1 if the worker's years of education (experience) are above the request from the recommended job.
3. Location match is 1 if the worker's city is consistent with the job's city.

Table 3.4: Set Difference Rate in Job Recommendations

| | Share |
|---|---|
| All | 0.0972 |
| Age | |
| Young | 0.1014 |
| Old | 0.0929 |
| Gender | |
| Female | 0.0939 |
| Neutral | 0.1023 |
| Male | 0.0938 |
| Hierarchy | |
| Entry | 0.0902 |
| Middle | 0.1029 |
| High | 0.0968 |
| City | |
| Beijing | 0.0961 |
| Shanghai | 0.1005 |
| Shenzhen | 0.0983 |
| Guangzhou | 0.0938 |
| | |
| Sample Size | 25,099 |

Notes:

1. Set difference rate is computed by the number of gender-specific jobs over the number of jobs recommended to both male and female applicants.

2. Duplicates of job recommendations from different rounds are counted once.

Table 3.5: Gender Differences on Explicit Measures of Recommended Jobs

| | Male – Female |
|---|---|
| Posted Wage | 2708.78* |
| | (1526.03) |
| Required Education | 0.0226 |
| | (0.0259) |
| Required Experience | 0.0779*** |
| | (0.0261) |

Notes:

1. Gender difference is computed from the mean of male-only jobs minus the mean of female-only jobs. Standard errors are in parentheses, which are derived from two-sample t-tests with equal variance. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 3.6: Gender Difference in Words in Job Recommendations

| | **Female Words** | **Male Words** |
|---|---|---|
| Skills | assist (-0.0195), data (-0.0195), administrative (-0.0188), speak (-0.0160), chat tools (-0.0122), documentation (-0.0101) | decision making (0.0071), design (0.0102), cooperation (0.0127), teamwork (0.141), engineering (0.0177), manage (0.0177), independent (0.0213) |
| Work Form | flexible (-0.0176), regular hour (-0.0124), weekly break (-0.0184) | commute (0.0129), work overtime (0.0202), nightwork (0.0084) |
| Benefits | base pay (-0.0200), marriage leave (-0.0091), maternity leave (-0.0089), social security (-0.0076), unemployment insurance (-0.0021), parental level (-0.0029) | shuttle (0.0139), medical insurance (0.0155), vacation (0.0168), meal (0.0213), reward (0.0256), stock (0.0269) |
| Company | training (-0.0133) | public company (0.0138) |
| Requirements | punctual (-0.0128), careful (-0.0125), outgoing (-0.0091), facial (-0.0026), figure (-0.0079), patient (-0.0068), healthy (-0.0064), temperament (0.0063), new grad (-0.0053), below 35 (-0.0050), non experience (-0.0041), trustworthy (-0.0032) | self-motivated (0.0073), pressure (0.0102), innovative (0.0107), experienced (0.0125), |

Notes:

1. Table 3.6 displays words that have significantly different probabilities of presenting in male-only and female-only jobs. Coefficients in parentheses represent the gender difference (male-female).
2. Female (male) words are defined from the proportion test with negative (positive) gender differences that are significant at 5% level.

Table 3.7: Gender Differences on Words and Gender Stereotypes

| | **Female Words** | **Male Words** |
|---|---|---|
| Skills | assist, data, administrative, speak, chat tools, documentation | decision making, design, cooperation, teamwork, engineering, manage, independent, leadership |
| Work Form | flexible, regular hour, weekly break | commute, work overtime, nightwork |
| Benefits | base pay, marriage leave, maternity leave, social security, unemployment insurance, parental level | shuttle, medical insurance, vacation, meal, reward, stock |
| Company | training | public company |
| Requirements | punctual, careful, outgoing, facial figure, patient, healthy, temperament, new grad, below 40, non experience, trustworthy | self-motivated, pressure, innovative, experienced |

Notes:

1. Table 3.7 shows the relation between gendered words in job ads and gender stereotypes. The color intensity indicates the maleness and femaleness consistency with gender stereotypes from literature and two survey results. Female words are highlighted with red colors, male words highlighted with blue colors, and strong color indicates high consistency.

Table 3.8: Top 10 Words in Prediction of Gender-Specific Recommended Jobs

| **OLS** | **Lasso** | **Ridge** | **Random Forest** |
|---|---|---|---|
| night work | base pay | night work | activities |
| trustworthy | regular hour | trustworthy | meal |
| below40 | marriage leave | work shift | commute |
| base pay | engineering | big and small week | holiday |
| regular hour | independent | facial | vacation |
| data | stock | marriage leave | commission |
| flexible | public | below40 | allowance |
| documentation | overtime | endowment ins | training |
| weekly break | flexible | maternity leave | fiveone |
| administrative | data | base pay | reward |

Notes:

1. Table 3.8 presents the top words in predicting whether a job is only recommended to male applicants. The outcome variable is binary and equals 1 for male-only jobs, and independent variables are 167 dummy variables for the existence of words in job ads.

2. Column 1 lists words from the OLS regression, which are significant at 5% level and sorted in descending order of the magnitude of coefficients.

3. Column 2 and 3 present words that are selected by the Lasso and Ridge regression. The penalty parameter for Lasso regression is 0.23 and is 0.31 in Ridge regression. Those are determined by using 20-fold cross-validation for the highest R squared. Words are sorted in descending order of the magnitude of estimation effects.

4. In column 4, random forest is applied to find words that have high impacts on the classification of male-only and female-only jobs based on 100 bootstraps and Gini impurity. Words are sorted in descending order of the importance factor.

5. "fiveone" represents "five social insurance and one housing fund" (五险一金), including endowment insurance, medical insurance, unemployment insurance, employment injury insurance, maternity insurance and housing fund. Big and small week describes the working schedule in which workers have one-day rest in one week and two-day rest in the next week.

Table 3.9: Effects of Views from Hiring Agents on Job Recommendations

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| ViewF | 0.0141* | 0.0188*** | 0.0165** | 0.0203*** |
| | (0.007) | (0.007) | (0.007) | (0.007) |
| ViewD | -0.0010 | -0.0008 | 0.0046 | 0.0072 |
| | (0.013) | (0.013) | (0.013) | (0.013) |
| Age | | Yes | Yes | Yes |
| Job Gender Type | | | Yes | Yes |
| Job Board | | | | Yes |
| N | 1033 | 1033 | 1033 | 1033 |
| $R^2$ | 0.0109 | 0.0212 | 0.0408 | 0.0633 |

Notes:

1. The dependent variable is the number of gender-specific jobs in 100 recommended jobs.
2. In column 1, the regressors are the number of views on the female's profile and the gender gap on the number of views (male-female) in each gender pair. Column 2 controls for young or older pairs. Column 3 further controls the worker's job gender type, including female-dominated jobs, gender-neutral jobs and male-dominated jobs. Column 4 adds the job board fixed effect.
3. Standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

# Appendix A

# Appendix for "Gender-Targeted Job Ads in the Recruitment Process: Facts from a Chinese Job Board"

## A.1 International Examples of Gendered Job Ads from Indeed.com

Accurate measures of the share of ads on a job site that are gender-targeted cannot be generated by conducting searches on the site's public portals, for a number of reasons. First, the ads on the site at a point in time represent a stock sample with potentially many stale ads. Second, unless the job board has chosen to collect and to publicize an unambiguous indicator of the employer's gender preference (as XMRC does), gender preferences can be expressed in many different ways, some of which are evasive, others of which are costly to detect.[1] Third, jobseekers' search results are often

[1]The use of gendered job titles (such as abogada and abogado in Spanish) is particularly burdensome to measure since each title expresses gender in a different way.

prioritized in ways that are opaque to the user. Finally, without a well-defined sample that has been drawn from the board's internal database, researchers are forced to rely on denominators provided by the job board, which are not clearly defined and prone to exaggeration.[2]

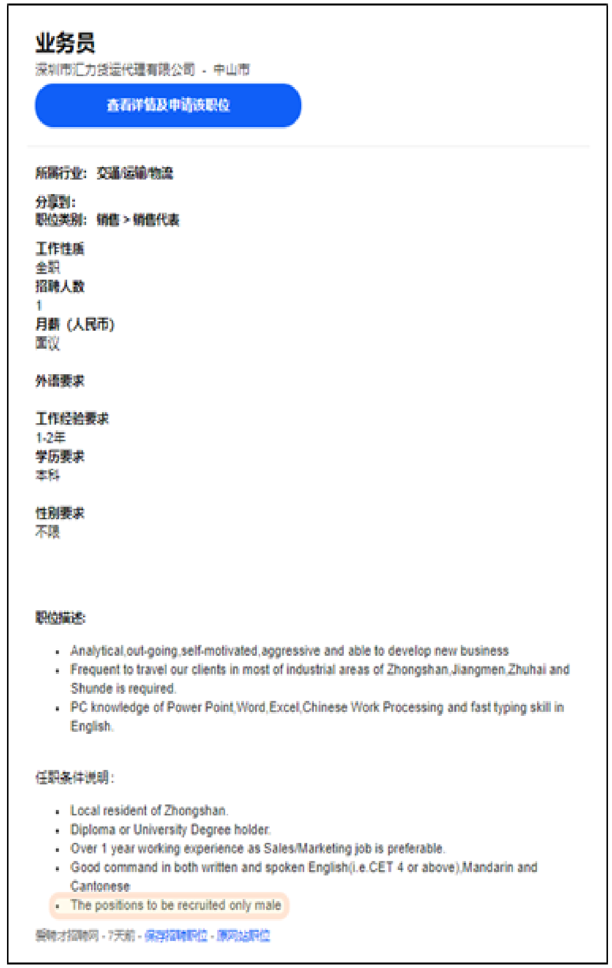
With these cautions in mind, we can arguably get some indicators of at least the presence and typical form of explicit gender requests by conducting keyword searches for jobs through the worker portal on a site. In this document we present examples of the results of such searches on Indeed.com, which currently operates job search platforms in 63 countries. The ads reproduced in the following pages were collected from Indeed.com's international portal: https://www.indeed.com/worldwide on November 12, 2018. In all cases, we searched for the terms "male" and "female" in the sites' native languages (this was English in India and Pakistan), then – where necessary – used Google to translate the results. Since "male" and "female" can be used in several ways that do not request a specific gender for the job (including saying that both men and women are welcome), we manually searched through these search results ads till we found ads that expressed a preference for one gender. We never had to go beyond the first 50 search results to find such ads. Noting that Indeed, as a U.S.-owned company, may be more sensitive to stigma associated with posting gendered ads, and that its international sites tend to serve educated and disproportionately English-speaking workers, it seems likely that gendered ads would be even easier to find on locally-owned and operated sites.

In all cases the searches were done without creating an account on Indeed, and without specifying a type of work or location—the only search term was "male" or

[2]Another emerging difficulty is the possibility that job boards are designing their worker-facing search algorithms to make certain forms of explicit gender requests hard to find via a keyword search, even though these requests are still present in ads (that are found via other keywords). We report some suggestive evidence of this in Section 3.2.

"female". No other filtering or ordering of results was done. The countries searched are the ten countries represented by Indeed with the largest populations. Since Indeed serves ten of the eleven largest countries, our results are for the world's 11 most populous countries with exception of Bangladesh, representing 57.4% of the world's population. The ads are numbered by country population rank.

## 1. China


业务员
深圳市汇力货运代理有限公司 · 中山市
查看详情及申请该职位

所属行业: 交通/运输/物流

分享到:
职位类别: 销售 > 销售代表

工作性质
全职
招聘人数
1
月薪（人民币）
面议

外语要求

工作经验要求
1-2年
学历要求
本科

性别要求
不限

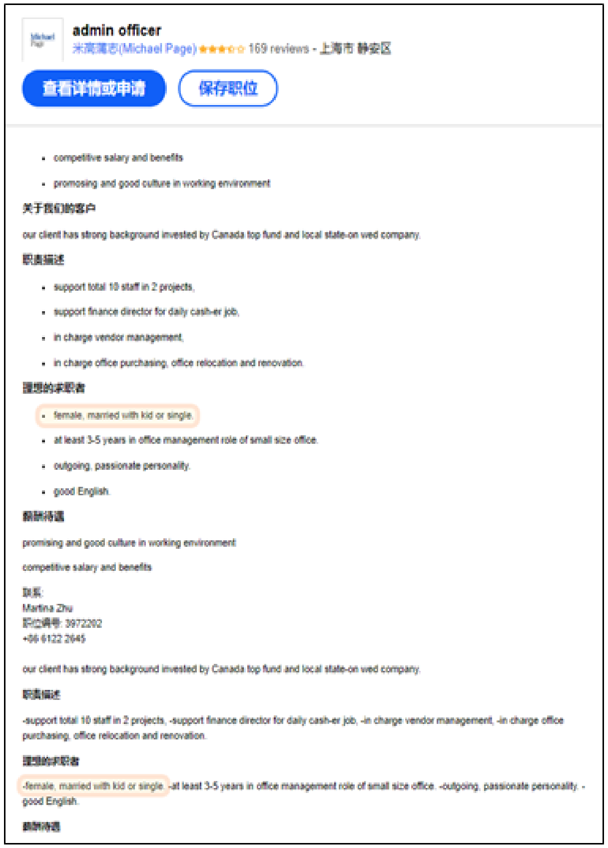
职位描述:

- Analytical,out-going,self-motivated,aggressive and able to develop new business
- Frequent to travel our clients in most of industrial areas of Zhongshan,Jiangmen,Zhuhai and Shunde is required.
- PC knowledge of Power Point,Word,Excel,Chinese Work Processing and fast typing skill in English.

任职条件说明:

- Local resident of Zhongshan.
- Diploma or University Degree holder.
- Over 1 year working experience as Sales/Marketing job is preferable.
- Good command in both written and spoken English(i.e.CET 4 or above),Mandarin and Cantonese
- The positions to be recruited only male

[illegible] - 7天前 - 保存招聘职位 - 原网站职位


admin officer
米高蒲志(Michael Page) 169 reviews - 上海市 静安区
查看详情或申请 保存职位

- competitive salary and benefits
- promosing and good culture in working environment

关于我们的客户

our client has strong background invested by Canada top fund and local state-on wed company.

职责描述

- support total 10 staff in 2 projects,
- support finance director for daily cash-er job,
- in charge vendor management,
- in charge office purchasing, office relocation and renovation.

理想的求职者

- female, married with kid or single.
- at least 3-5 years in office management role of small size office.
- outgoing, passionate personality.
- good English.

薪酬待遇

promising and good culture in working environment

competitive salary and benefits

联系:
Martina Zhu
职位编号: 3972202
+86 6122 2645

our client has strong background invested by Canada top fund and local state-on wed company.

职责描述

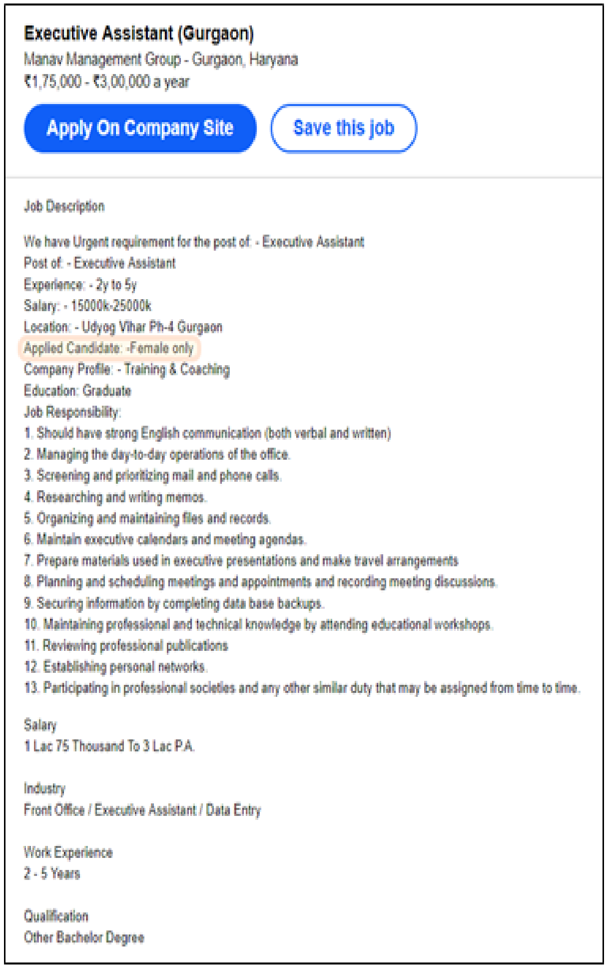
-support total 10 staff in 2 projects, -support finance director for daily cash-er job, -in charge vendor management, -in charge office purchasing, office relocation and renovation.

理想的求职者

-female, married with kid or single. -at least 3-5 years in office management role of small size office. -outgoing, passionate personality. -good English.

薪酬待遇

## 2. India



**Executive Assistant (Gurgaon)**
Manav Management Group - Gurgaon, Haryana
₹1,75,000 - ₹3,00,000 a year

Apply On Company Site | Save this job

Job Description

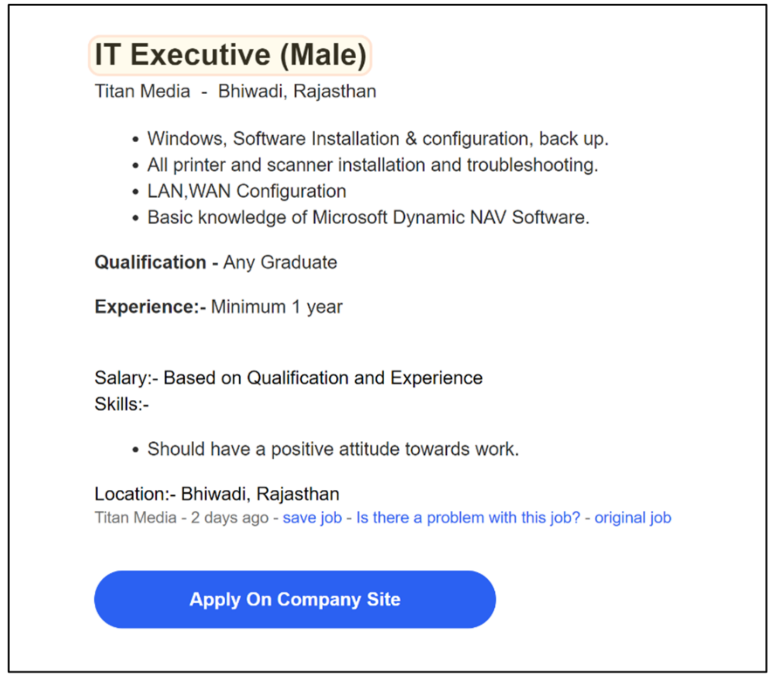
We have Urgent requirement for the post of: - Executive Assistant
Post of: - Executive Assistant
Experience: - 2y to 5y
Salary: - 15000k-25000k
Location: - Udyog Vihar Ph-4 Gurgaon
Applied Candidate: -Female only
Company Profile: - Training & Coaching
Education: Graduate
Job Responsibility:
1. Should have strong English communication (both verbal and written)
2. Managing the day-to-day operations of the office.
3. Screening and prioritizing mail and phone calls.
4. Researching and writing memos.
5. Organizing and maintaining files and records.
6. Maintain executive calendars and meeting agendas.
7. Prepare materials used in executive presentations and make travel arrangements
8. Planning and scheduling meetings and appointments and recording meeting discussions.
9. Securing information by completing data base backups.
10. Maintaining professional and technical knowledge by attending educational workshops.
11. Reviewing professional publications
12. Establishing personal networks.
13. Participating in professional societies and any other similar duty that may be assigned from time to time.

Salary
1 Lac 75 Thousand To 3 Lac P.A.

Industry
Front Office / Executive Assistant / Data Entry

Work Experience
2 - 5 Years

Qualification
Other Bachelor Degree



**IT Executive (Male)**
Titan Media - Bhiwadi, Rajasthan

- Windows, Software Installation & configuration, back up.
- All printer and scanner installation and troubleshooting.
- LAN,WAN Configuration
- Basic knowledge of Microsoft Dynamic NAV Software.

**Qualification -** Any Graduate

**Experience:-** Minimum 1 year

Salary:- Based on Qualification and Experience
Skills:-

- Should have a positive attitude towards work.

Location:- Bhiwadi, Rajasthan
Titan Media - 2 days ago - save job - Is there a problem with this job? - original job

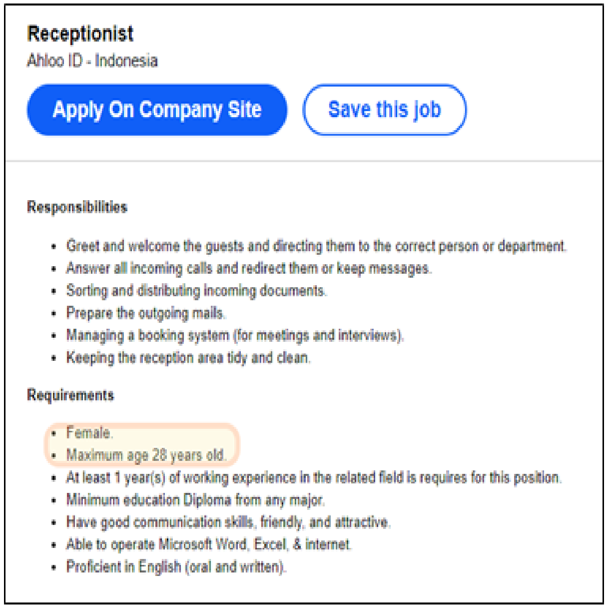
Apply On Company Site

**3. USA**

| Female | Male |
| --- | --- |
| All of the **first 50 hits for "women"** were used to convey:<br>• a "genuine" job requirement (e.g. customer service swimwear, TSA pat-down officer)<br>• a feature of the work environment (e.g. "female run and managed company", support staff for female clients in drug recovery)<br>• a diversity statement (e.g. "EOE/Minorities/Females/Veterans/Disabled", in 39/50 ads)<br>• different physical qualifications for men and women (e.g. "Correctional officer … 4 pushups female, 8 pushups male")<br>with one possible exception: "front desk agent--we are looking to add another female to our front desk position… professional appearance". | All of the **first 50 hits for "men"** were used to convey:<br>• a "genuine" job requirement (e.g. housekeeper for a men's locker room, male clothing model, male urine sample collection specialist)<br>• a feature of the work environment (e.g. hairstylist for male clientele, clerk in male inmate facility), or<br>• a military draft requirement (e.g. "Census enumerator--all male applicants must be registered with Selective Service system".<br>with one possible exception: "male on-camera sports host". |

We searched for "male" and "female" as keywords without registering as workers and without specifying a location or type of worker.

### 4. Indonesia



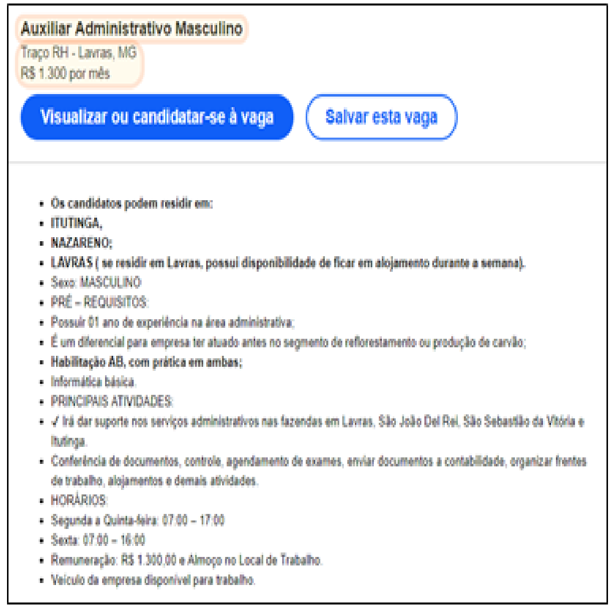


In several countries, requests for a specific gender are frequently accompanied by a desired age range as well. See Helleseter et al. (2020) for detailed evidence on age*gender interactions in job ads from China and Mexico.

### 5. Brazil

Auxiliar Administrativo Feminino em Lavras
Traço RH - Lavras, MG
R$ 1.000 por mês
Visualizar ou candidatar-se à vaga
Salvar esta vaga

- PRÉ – REQUISITO
- ✓ Possuir 01 ano de experiência na área administrativa;
- ✓ Conhecimento em informática e boa digitação.
- PRINCIPAIS ATIVIDADES:
- ✓ Atuar com atendimento ao cliente, confecção de contratos, rescisões, atendimento telefônico, resolver pendência sobre imóveis e demais atividades.
- ✅ HORÁRIO:
- Segunda a Sexta: 08:30 – 18:00 hs.
- Sábados eventuais.
- Remuneração: R$ 1.000,00



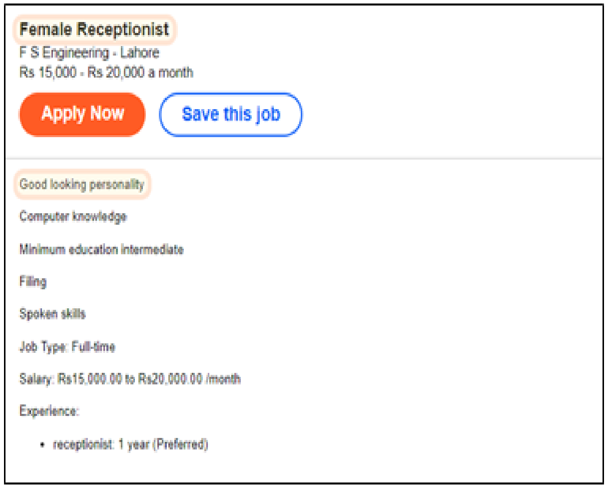
Auxiliar Administrativo Masculino
Traço RH - Lavras, MG
R$ 1.300 por mês
Visualizar ou candidatar-se à vaga
Salvar esta vaga

- **Os candidatos podem residir em:**
- **ITUTINGA,**
- **NAZARENO;**
- **LAVRAS ( se residir em Lavras, possui disponibilidade de ficar em alojamento durante a semana).**
- Sexo: MASCULINO
- PRÉ – REQUISITOS:
- Possuir 01 ano de experiência na área administrativa;
- É um diferencial para empresa ter atuado antes no segmento de reflorestamento ou produção de carvão;
- **Habilitação AB, com prática em ambas;**
- Informática básica.
- PRINCIPAIS ATIVIDADES:
- ✓ Irá dar suporte nos serviços administrativos nas fazendas em Lavras, São João Del Rei, São Sebastião da Vitória e Itutinga.
- Conferência de documentos, controle, agendamento de exames, enviar documentos a contabilidade, organizar frentes de trabalho, alojamentos e demais atividades.
- HORÁRIOS:
- Segunda a Quinta-feira: 07:00 – 17:00
- Sexta: 07:00 – 16:00
- Remuneração: R$ 1.300,00 e Almoço no Local de Trabalho.
- Veículo da empresa disponível para trabalho.

A large number of Indeed's Brazilian ads say the job is open to both men and women, but single-sex ads like these also exist.

This is an interesting example of the same company is advertising similar jobs for men and women, but offering a 30 percent higher wage in the male ad.

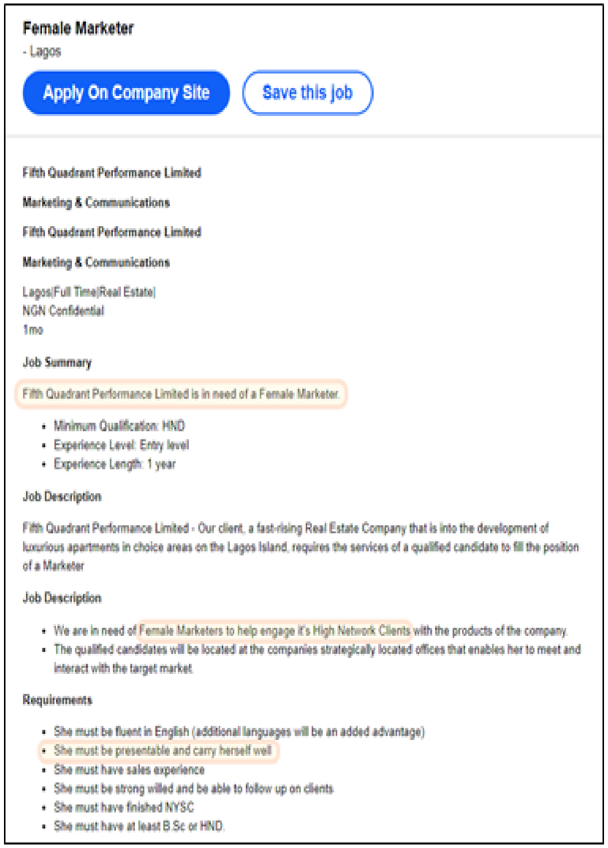
### 6. Pakistan





Requests for women in customer-contact jobs like this one frequently include explicit requests for beauty (Helleseter et al., 2020).

## 7. Nigeria



**Female Marketer**

- Lagos

Apply On Company Site | Save this job

**Fifth Quadrant Performance Limited**

**Marketing & Communications**

**Fifth Quadrant Performance Limited**

**Marketing & Communications**

Lagos|Full Time|Real Estate|
NGN Confidential
1mo

**Job Summary**

Fifth Quadrant Performance Limited is in need of a Female Marketer.

- Minimum Qualification: HND
- Experience Level: Entry level
- Experience Length: 1 year

**Job Description**

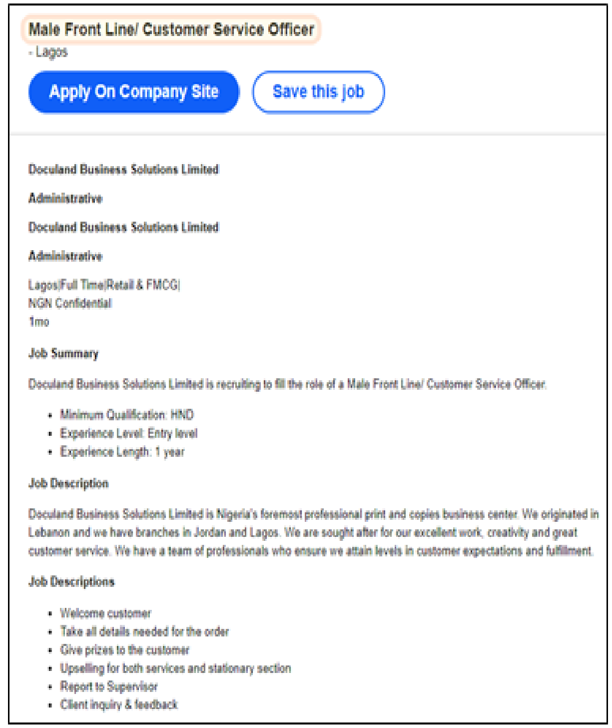
Fifth Quadrant Performance Limited - Our client, a fast-rising Real Estate Company that is into the development of luxurious apartments in choice areas on the Lagos Island, requires the services of a qualified candidate to fill the position of a Marketer

**Job Description**

- We are in need of Female Marketers to help engage it's High Network Clients with the products of the company.
- The qualified candidates will be located at the companies strategically located offices that enables her to meet and interact with the target market

**Requirements**

- She must be fluent in English (additional languages will be an added advantage)
- She must be presentable and carry herself well
- She must have sales experience
- She must be strong willed and be able to follow up on clients
- She must have finished NYSC
- She must have at least B.Sc or HND.



**Male Front Line/ Customer Service Officer**

- Lagos

Apply On Company Site | Save this job

**Doculand Business Solutions Limited**

**Administrative**

**Doculand Business Solutions Limited**

**Administrative**

Lagos|Full Time|Retail & FMCG|
NGN Confidential
1mo

**Job Summary**

Doculand Business Solutions Limited is recruiting to fill the role of a Male Front Line/ Customer Service Officer.

- Minimum Qualification: HND
- Experience Level: Entry level
- Experience Length: 1 year

**Job Description**

Doculand Business Solutions Limited is Nigeria's foremost professional print and copies business center. We originated in Lebanon and we have branches in Jordan and Lagos. We are sought after for our excellent work, creativity and great customer service. We have a team of professionals who ensure we attain levels in customer expectations and fulfillment.

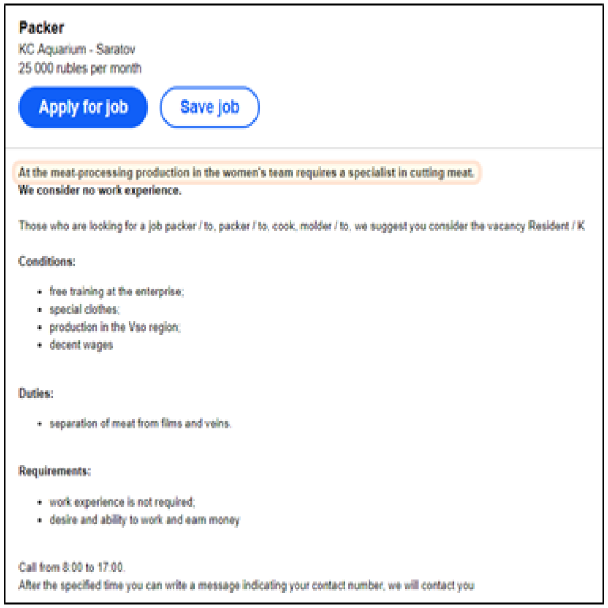
**Job Descriptions**

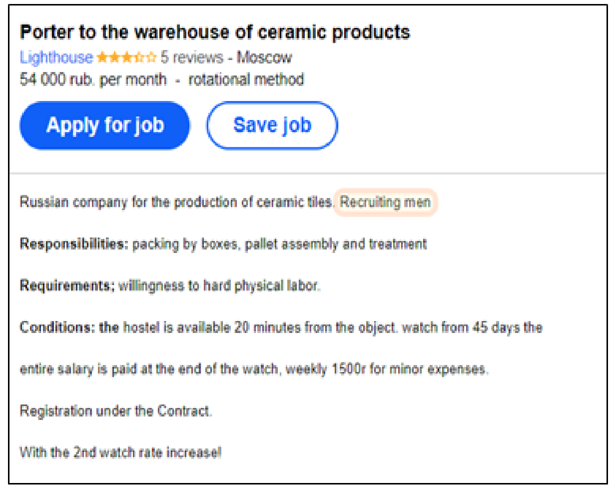
- Welcome customer
- Take all details needed for the order
- Give prizes to the customer
- Upselling for both services and stationary section
- Report to Supervisor
- Client inquiry & feedback

## 8. Bangladesh

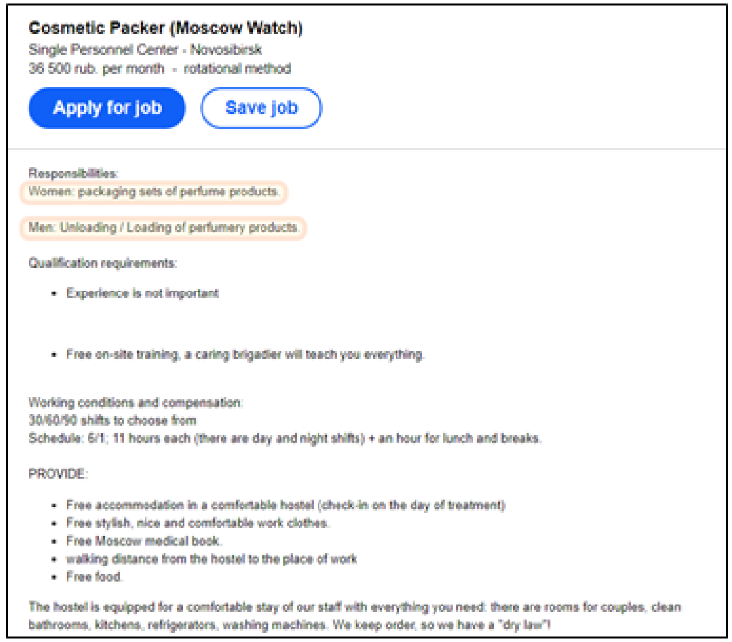
Does not have an Indeed site.

## 9. Russia



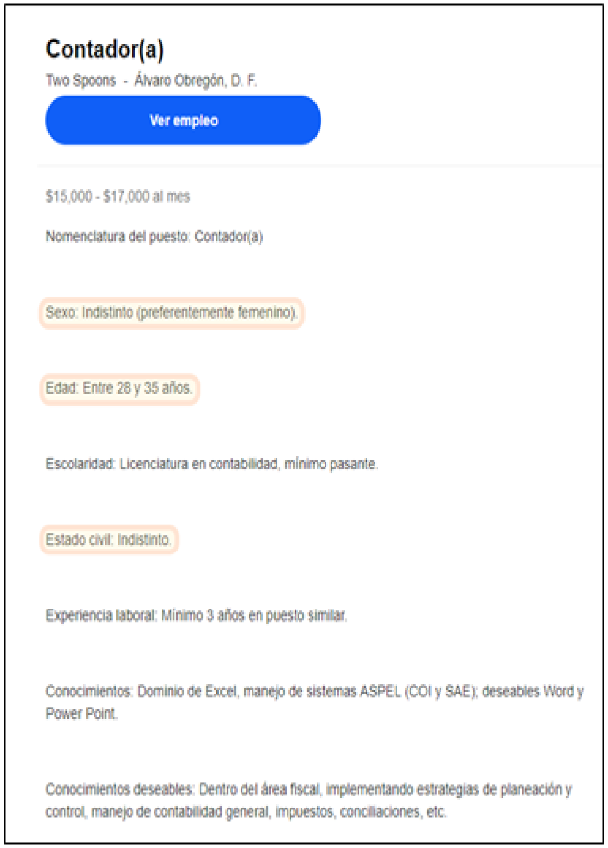



**9. Russia (continued)**




Ads of this type – where a company requests both men and women, but for different duties within the firm – were much more common on Indeed's Russia site than ads requesting a single gender only.

## 10. Mexico


Contador(a)
Two Spoons - Álvaro Obregón, D. F.
Ver empleo

$15,000 - $17,000 al mes

Nomenclatura del puesto: Contador(a)

Sexo: Indistinto (preferentemente femenino).

Edad: Entre 28 y 35 años.

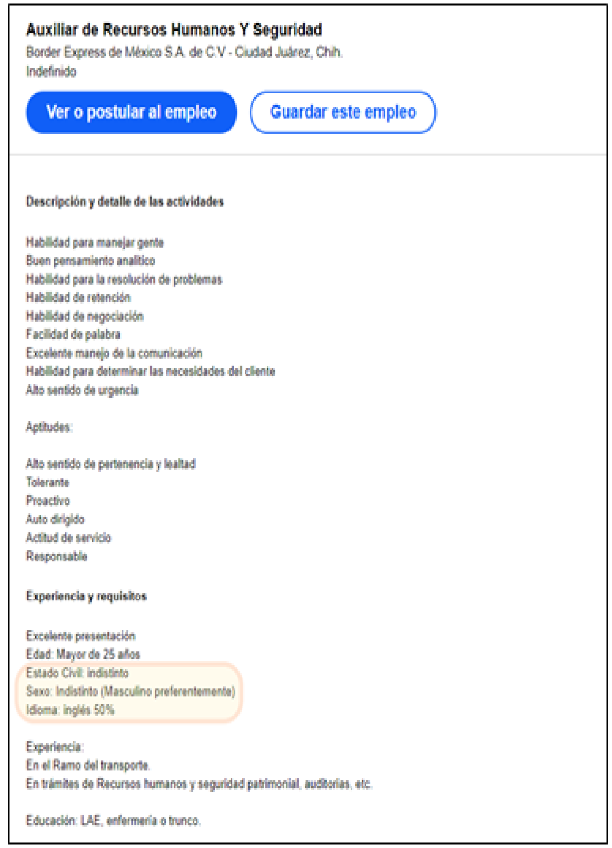
Escolaridad: Licenciatura en contabilidad, mínimo pasante.

Estado civil: Indistinto.

Experiencia laboral: Mínimo 3 años en puesto similar.

Conocimientos: Dominio de Excel, manejo de sistemas ASPEL (COI y SAE); deseables Word y Power Point.

Conocimientos deseables: Dentro del área fiscal, implementando estrategias de planeación y control, manejo de contabilidad general, impuestos, conciliaciones, etc.


Auxiliar de Recursos Humanos Y Seguridad
Border Express de México S.A. de C.V - Ciudad Juárez, Chih.
Indefinido
Ver o postular al empleo
Guardar este empleo

Descripción y detalle de las actividades

Habilidad para manejar gente
Buen pensamiento analítico
Habilidad para la resolución de problemas
Habilidad de retención
Habilidad de negociación
Facilidad de palabra
Excelente manejo de la comunicación
Habilidad para determinar las necesidades del cliente
Alto sentido de urgencia

Aptitudes:

Alto sentido de pertenencia y lealtad
Tolerante
Proactivo
Auto dirigido
Actitud de servicio
Responsable

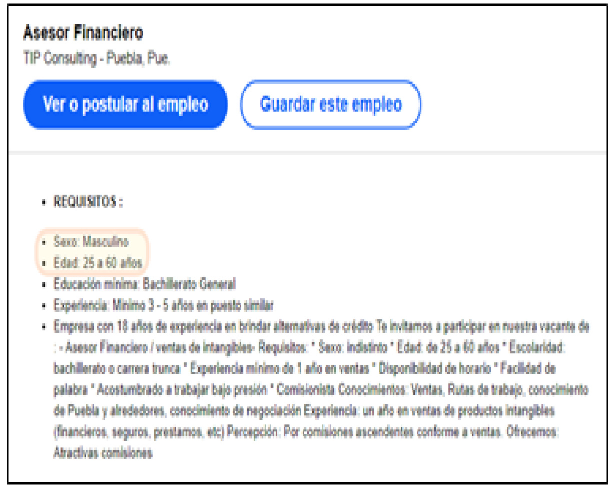
Experiencia y requisitos

Excelente presentación
Edad: Mayor de 25 años
Estado Civil: indistinto
Sexo: Indistinto (Masculino preferentemente)
Idioma: inglés 50%

Experiencia:
En el Ramo del transporte.
En trámites de Recursos humanos y seguridad patrimonial, auditorías, etc.

Educación: LAE, enfermería o trunco.

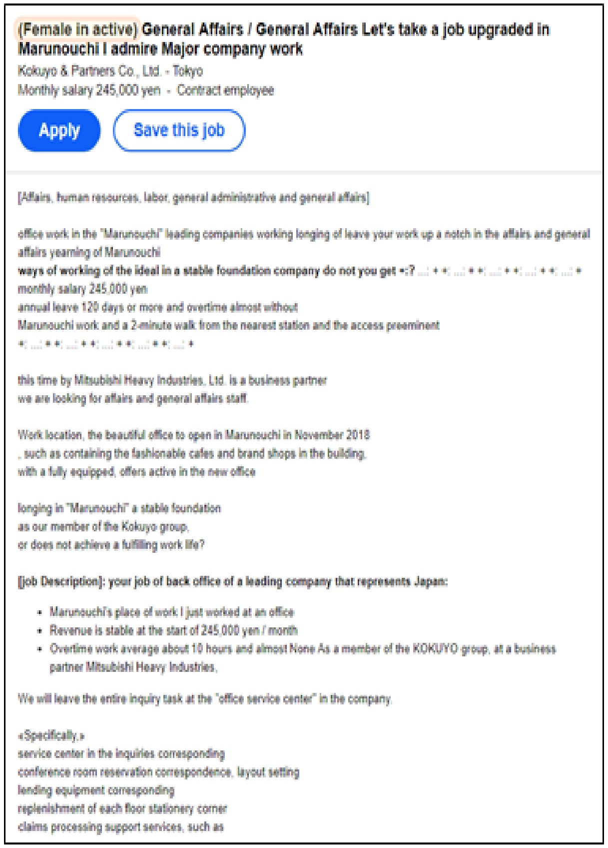
## 10. Mexico(continued)

### 11. Japan

The Japanese Equal Employment Opportunity Law prohibits employers from saying that they prefer to hire men (women). However, job ads on Indeed's Japan site frequently say that men (women) are playing "active roles", or 'thriving' in these jobs or in the firm. The intent appears to be to signal that the jobs in question are suited to a particular gender.

Female



(Female in active) **General Affairs / General Affairs Let's take a job upgraded in Marunouchi I admire Major company work**

Kokuyo & Partners Co., Ltd. - Tokyo

Monthly salary 245,000 yen - Contract employee

Apply

Save this job

[Affairs, human resources, labor, general administrative and general affairs]

office work in the "Marunouchi" leading companies working longing of leave your work up a notch in the affairs and general affairs yearning of Marunouchi

**ways of working of the ideal in a stable foundation company do not you get**

monthly salary 245,000 yen

annual leave 120 days or more and overtime almost without

Marunouchi work and a 2-minute walk from the nearest station and the access preeminent

this time by Mitsubishi Heavy Industries, Ltd. is a business partner

we are looking for affairs and general affairs staff.

Work location, the beautiful office to open in Marunouchi in November 2018

, such as containing the fashionable cafes and brand shops in the building,

with a fully equipped, offers active in the new office

longing in "Marunouchi" a stable foundation

as our member of the Kokuyo group,

or does not achieve a fulfilling work life?

**[job Description]: your job of back office of a leading company that represents Japan:**

- Marunouchi's place of work I just worked at an office
- Revenue is stable at the start of 245,000 yen / month
- Overtime work average about 10 hours and almost None As a member of the KOKUYO group, at a business partner Mitsubishi Heavy Industries,

We will leave the entire inquiry task at the "office service center" in the company.

«Specifically,»

service center in the inquiries corresponding

conference room reservation correspondence, layout setting

lending equipment corresponding

replenishment of each floor stationery corner

claims processing support services, such as

(女性活躍中)**総務・庶務 憧れの丸の内でワンランク上のお仕事をお任せ 大手企業勤務**

コクヨ&パートナーズ株式会社 - 東京都

月給 24.5万円 - 契約社員

応募する

この求人を保存する

【総務、人事、労務、一般事務・庶務】

総務・庶務 憧れの丸の内でワンランク上のお仕事をお任せ 大手企業勤務 憧れの「丸の内」でのオフィスワーク

**安定基盤のある会社で理想の働き方を手に入れませんか？**

月給24万5000円~

年休120日以上&残業ほぼなし

丸の内勤務&最寄駅から徒歩2分とアクセス抜群

今回はお取引先である三菱重工業株式会社にて

総務・庶務スタッフを募集しています。

勤務地は、2018年11月に丸の内にオープンするキレイなオフィス

ビル内にはおしゃれなカフェやブランドショップが入っているなど、

充実した設備を持つ、新オフィスでご活躍いただけます

憧れの「丸の内」で安定基盤のある

コクヨグループの当社の一員として、

充実したワークライフを実現しませんか？

**【仕事内容】：日本を代表する大手企業でのバックオフィスのお仕事：**

- 勤務地は丸の内 OPENしたばかりオフィスで働ける
- 月給24万5000円スタートで収入も安定
- 残業月平均10時間程度とほぼなし コクヨグループの一員として、取引先である三菱重工業株式会社にて、

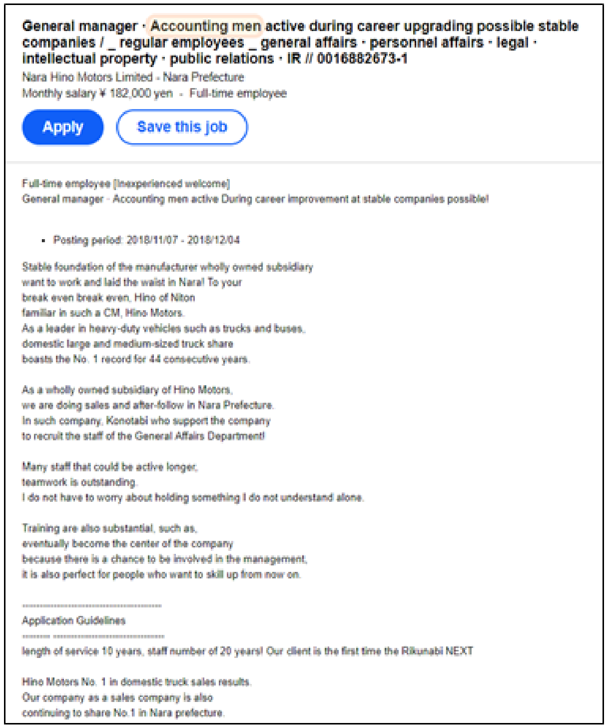
社内の「オフィスサービスセンター」での問い合わせ業務全般をお任せします。

＜具体的には＞

サービスセンターでのお問い合わせ対応

A better translation of the job title is: "(Women thriving) General Affairs. Asked to perform a step-up task in the enviable Marunouchi area. Working for major corporations." This is a job ad for a contract firm. The successful applicant would work in a new office of this contract firm in a new building in Marunouchi area, and work for one of its clients, Mitsubishi Heavy Industry. Since Japanese employers are not allowed to make explicit gender requests, "women thriving" is a way to signal that women are doing well in this particular position.

Male



**General manager · Accounting men active during career upgrading possible stable companies / _ regular employees _ general affairs · personnel affairs · legal · intellectual property · public relations · IR // 0016882673-1**

Nara Hino Motors Limited - Nara Prefecture
Monthly salary ¥ 182,000 yen - Full-time employee

Apply
Save this job

Full-time employee [Inexperienced welcome]
General manager - Accounting men active During career improvement at stable companies possible!

- Posting period: 2018/11/07 - 2018/12/04

Stable foundation of the manufacturer wholly owned subsidiary
want to work and laid the waist in Nara! To your
break even break even. Hino of Niton
familiar in such a CM, Hino Motors.
As a leader in heavy-duty vehicles such as trucks and buses,
domestic large and medium-sized truck share
boasts the No. 1 record for 44 consecutive years.

As a wholly owned subsidiary of Hino Motors,
we are doing sales and after-follow in Nara Prefecture.
In such company, Konotabi who support the company
to recruit the staff of the General Affairs Department!

Many staff that could be active longer,
teamwork is outstanding.
I do not have to worry about holding something I do not understand alone.

Training are also substantial, such as,
eventually become the center of the company
because there is a chance to be involved in the management,
it is also perfect for people who want to skill up from now on.

Application Guidelines

length of service 10 years, staff number of 20 years! Our client is the first time the Rikunabi NEXT

Hino Motors No. 1 in domestic truck sales results.
Our company as a sales company is also
continuing to share No.1 in Nara prefecture.

**総務・経理男性活躍中安定企業でキャリアアップ可能/_正社員_総務・人事・法務・知財・広報・IR//0016882673-1**

奈良日野自動車株式会社 - 奈良県
月給 18.2万円 - 正社員

応募する
この求人を保存する

正社員[未経験歓迎]
総務・経理 男性活躍中 安定企業でキャリアアップ可能!

- 掲載期間:2018/11/07 ~ 2018/12/04

メーカー100%子会社の安定基盤
奈良で腰をすえて働きたいあなたへ!
トントントントン、日野の2トン
そんなCMでおなじみの、日野自動車。
トラックやバスなど大型車両のリーディングカンパニーとして、
国内大・中型トラックシェアは
44年連続No.1の実績を誇っています。

日野自動車の100%子会社として、
奈良県内での販売やアフターフォローをしている私たち。
そんな当社で、このたび社内を支えてくださる
総務部のスタッフを募集します!

長く活躍してくれているスタッフが多く、
チームワークは抜群。
わからないことを一人で抱え込む心配はありません。

研修なども充実しており、
ゆくゆくは当社の中枢となって
経営にも携われるチャンスがあるので、
これからスキルアップしたい人にもピッタリですよ。

A better translation of the job title is: "General Affairs or Accounting. Men thriving. Possible to advance your career in a stable company."

This Appendix was prepared with the assistance of Steve Li and Jia You, undergraduate students at UCSB. The authors thank Takao Kato, Professor of Economics, Colgate University for helping me understand the Japanese ads.

## A.2 Legislation Affecting Gender-Targeted Job Ads in China

### A.2.1 Early Laws and Regulations Concerning Gender Discrimination

China's constitution and labor law have prohibited gender discrimination since at least 1982. For example, Article 48 of the Constitution of the People's Republic of China (1982) grants women "equal rights with men in all spheres of life, political, economic, cultural, social, and family life", and affirms the principle of equal pay for equal work for men and women. With the exception of "types of work that are not suitable for females", the Labor Law of the PRC (1994; Article 13) prohibits using sex as a pretext for excluding females from employment or for raising recruitment standards; similar provisions are found in the Law of the PRC on the Protection of Rights and Interests of Women (2005; Article 22), and the Law of the PRC on Promotion of Employment (2007, Articles 26 and 27.) The latter law also prohibits employment contracts that restrict female workers from getting married or bearing a child.

While a ban on ads (of any kind) that "carry any nationality, religious or sex discriminating information" has been in place since 1994 (Advertisement law of the PRC, Articles 7 and 39), the earliest regulations we are aware of that specifically prohibit gender discrimination by labor market intermediaries date from 2007. At that time, the Ministry of Labor and Social Security's Regulations on Employment Service and Employment Management prohibited intermediaries from "releasing any information indicating employment discrimination" (Articles 58 and 74).

Enforcement of China's anti-discrimination laws before 2012 however, is widely perceived to have been weak (Watch/Asia, 2018), and our previous studies of online job boards (Kuhn and Shen, 2013; Helleseter et al., 2020) suggest that these laws did not seriously constrain employers' use of explicitly gendered job ads at that time.

### A.2.2 Court Cases

According to Flory et al. (2015), the first lawsuit claiming gender discrimination in China's labor market was filed in July 2012. After graduating from a Beijing university, Ju Cao was told that she was not qualified for an administrative assistant job because "this was a position for men, we would not consider you although you are qualified". As part of an out-of-court settlement, the firm made a public apology to Ms. Cao. In 2014, another new graduate, Guo Mou was rejected from a copywriting job at Hangzhou's prestigious New East Cuisine Education school, for the reason that "men are more qualified for this position". The school was ordered to pay Ms. Guo 2,000 yuan for "spiritual injury" (CCTV.com, 2015). In China's first lawsuit on gender discrimination against a state-owned enterprise (SOE), Hu Ma was rejected for a delivery job with China Post. In response to her lawsuit, submitted on January 26,. 2015, China Post argued that delivery required workers to hold heavy objects, which met the legal exception of not being "suitable for females". The Court of Beijing rejected China Post's argument and ordered them to compensate Ms. Hu (Zhang, 2016).

Since the latter two lawsuits, the plaintiffs (Guo and Hu), have become activists against gender discrimination in employment. As part of their efforts, they have collected gender-targeted job ads on sites including Zhaopin.com, 51job.com, 58.com, Chinahr.com, and reported them to Ministry of Labor and Social Security.

In addition to the above court cases, a recent regulatory development seems to have prodded China's largest job boards to actively discourage and remove gendered job ads from their sites. In May 2016, China's Ministry of Industry and Information Technology issued a regulation aimed directly at gendered job ads on online job platforms. A key component of this regulation clarified the division of fines between the job board (30%) and the firm placing the ad (70%). This appears to have been at least partially effective: by October 2018, explicit requests for men or women ads were effectively absent from the two of largest privately operated job boards: 51 job and Zhaopin (see Appendix A.3).

Some insight into how this change occurred is available from our conversations with officials at Liepin.com, a 'high-end' job board catering to executive-level positions. After receiving notice of the May 2016 regulation, Liepin sent a letter to all HR personnel using their website, stating that the HR personnel would not be allowed to post new job ads stating an explicit preference for one gender. Hiring managers were also asked to revise existing ads by removing any gender labels or other statements of gender preference.

At the same time, Liepin developed and improved its own filtering system to detect gendered job ads. Focusing first on newly-posted ads, Liepin tagged ads including statements like "male first", or "only for women" "male engineer" etc. and asked HR personnel to change these ads. Starting in July 2016, Liepin actively revised previously-posted ads by removing the gender requests without changing anything else. All such ads were replaced by the end of August, 2016. Since then, in part due to increased scrutiny from applicants who are willing to report violations to the government, Liepin has improved its screening for words that may convey a preferred gender, using human screeners to examine jobs that are considered suspect by Liepin's algorithms. Notably, throughout this process, Liepin continued to allow HR personnel to filter job applications by gender, so that the firm could choose to see only applications from one gender regardless of who applied. Thus, at least on Liepin, internal filters seem to have replaced public gender requests.

## Appendix A.2 References

CCTV.com. 2015. “The Communist Party announced the top ten cases of labor violations: the first case of employment gender discrimination ranked first” CCTV News, February 02, 2015 (accessed November 18, 2018).

FlorCruz, Michelle. 2014. “Chinese Woman Wins Settlement In China's First Ever Gender Discrimination Lawsuit” International Business Times, February 2, 2014 (accessed November 18, 2018).

Zhang, Yuanyuan. 2016.“Ma Hu v. Beijing Post Employment Sex Discrimination Case Second Trial” China Women's News, January 29, 2016 (accessed November 18, 2018).

## A.3 Gendered Job Ads in China, 2018

### A.3.1 Methods

As noted in Appendix A.1, accurate measures of the share of ads on a job site that are gender-targeted cannot be generated by conducting searches on the site's public portals. As we did in Appendix A.1 for the international context, however, this Appendix attempts to document the presence and typical form of gender requests on various Chinese job sites by searching for jobs using gender-related keywords. Specifically, entering the sites via their jobseeker portals, we searched for words that might convey a gender preference by the employer. Then, we inspected the first page of results (usually 50 ads) to count the number of those hits in which the keyword was used to request a specific gender (as opposed to describing the product/service, or inviting both genders to apply). In performing these searches, we did not create a worker profile on the site or specify any worker characteristics, nor did we enter any search terms for the location and type of work sought. All searches were performed in October 2018. The only search terms we entered were the following (one at a time):

1. Direct gender indicators: "man (男)" and "woman (女)" (This includes "men" and "women" in Chinese).

2. Transformed gender indicators: "nan" (the pronunciation for man, or 南 meaning south, which has the same pronunciation with man in Chinese), and "nv" (the pronunciation for woman; nv is the Chinese phonetic of women). These indicators have been used by employers to evade some recent enforcement activities (Watch/Asia, 2018).

3. Gendered adjectives: handsome (帅), gentleman (绅士), and "tall and strong" (高大健壮) for men; beautiful (美丽), lady (淑女), and "beautiful face" (面容姣

好) for women.

4. New "web words": little brother (小哥 lad), little sister (小姐姐 lass). These new words refer in a polite way to someone who is young and good-looking. They are more widely used by young people, and in job ads aimed at younger workers, such as social media jobs.

In the rest of this Appendix, we provide a verbal overview of these search results. Tabular results with additional details and commentary are available from the authors.

### A.3.2 51job.com and Zhaopin.com

51job and Zhaopin are China's two largest job boards. Both are privately run and cater to private-sector firms and workers. Our searches of these sites revealed no uses of the 'direct' indicators "man (男)" and "woman (女)" to request a specific gender, and only a few uses of the transformed indicators "nan", or "nv". One likely reason is enforcement: these boards now face a risk of being fined if they post gendered jobs; in response, the boards seem to have improved the screening of sensitive words so they no longer appear in workers' search results. In addition, these boards now discourage recruiters from making gender requests in job ads.[3] A second possible reason is that these boards cater to highly skilled workers; this may leave the boards more vulnerable to disapproval on social media if they post gendered ads. A third contributing factor may be the fact that employers' demand for gender profiling was relatively low in highly skilled jobs to begin with, even when this practice was widely tolerated (Kuhn and Shen, 2013; Helleseter et al., 2020), thus reducing the cost of compliance with the

[3] Zhaopin's portal states "Please do not include words that have the meaning of gender discrimination". Chinahr says "To make sure the job ad can pass checking, please do not enter repeat or meaningless information, and do not enter discriminating information, such as 'women first', or 'only for men'".

new restrictions.

This being noted, our analysis also shows that these job boards still accept subtler gender signals in ads, such as the gendered adjectives and the new “web words” we examined. For example, even though searches for “woman” yielded no results, searches for compound words like “lady” = “gentle+woman” (two characters) yielded several pages of results (though most of these refer to the names, products or brand of the firms). In addition, the adjectives “handsome”, “gentleman”, and “tall and strong” were frequently used to request men in jobs that included fitness instructors, sales, and warehouse work. “Beautiful”, “lady”, and “beautiful face” were used to request women in jobs that included customer service, front desk and modeling. Finally, the new web words ”little brother” and “little sister” were also used to convey a clear gender preference. For example “little brother” was frequently used to request (young) men for (electric bicycle) delivery jobs, and “little sister” for camgirl jobs.[4] Also of interest, both Zhaopin and 51job allow recruiters to select a filter that will only show the recruiter the applications from a particular gender.[5] Overall, prohibition of gendered job ads has pushed formerly overt discrimination into more hidden forms on these platforms.

### A.3.3 Chinahr.com

We conducted a comparable search of Chinahr.com, a national job board that caters more to blue collar workers than Zhaopin and 51job. Here, the terms “man” and “woman’ each yielded more than one page of search results.[6] Inspecting the first page

[4]The delivery jobs in question involve driving electric bicycles with packages or meals; pay is commission-based and the jobs are short term and relatively dangerous. Because most of the employees are young men, they are typically called “delivery little brother”.

[5]The same is true for Liepin.com, a recruiting site focusing on higher managerial positions.

[6]On Chinahr, a page of search results comprises 40 job ads.

of these revealed that 17 (or 43%) of the uses of “man” were explicit gender requests, as were 15 (or 38%) of the uses of “woman”. Interestingly, here the transformed gender terms “nan” and “nv” were almost never used to request an applicant gender, perhaps because direct requests were still feasible. Perhaps for the same reason, gendered adjectives and new web words – while present – weren't used much to request candidates of a specific gender either. We speculate that Chinahr is more tolerant of gender profiling by employers than 51job and Zhaopin because of its focus on blue collar jobs, where, as noted, employers' demand for gender profiling appears to be much higher (Kuhn and Shen, 2013; Helleseter et al., 2020), and where both stigma and enforcement may be weaker.

### A.3.4 Local Internet Job Boards

Parallel to the private-sector boards discussed above, China has a system of government-run or government-sponsored job boards that operate at the city or province level. These boards' names end in RC, GGZP or HR; XMRC is one of them. In general, these boards tend to serve lower skill levels than the national boards described previously. Like the national job boards, however, all of these boards serve private-sector employers and workers; recruiting for government jobs takes place via other channels. In a comprehensive web search – also in November 2018 – we were able to find 33 such boards of non-negligible size.[7]

When we examined the recruiter portals of these 33 sites, we found that 11 of them (including XMRC) asked employers to specify the gender of the worker they were seeking when the employer fills out a template for a job ad. Four of the sites (also including XMRC) allowed workers to filter job ads based on these employer requests. Keyword

[7] We found 57 boards in total, but 24 of these claimed to host 1000 or fewer job ads.

searches for "male" and "female" produced hits on all but two of these sites, and examination of the first 50 hits on each site revealed that these terms were frequently used to express a preference for male or female applicants. Code words like "nan" and "nv" turned up almost no results, perhaps because direct gender requests are still possible on these sites.

In sum, compared with private job boards, government-sponsored local job sites had a larger number of explicitly gendered job ads in late 2018. We can think of three possible reasons for this. First, these sites tend to be relatively small, so they may so far have escaped the attention of regulators. Second, these sites – especially the pure job-posting services – serve less-skilled jobs and workers, where employers' demand for gender filters is considerably greater. Finally, in China, workers may be much less inclined to report government-sponsored sites for regulatory violations, compared to privately operated sites. Since November 2018, increasing enforcement appears, however, to have encroached on these job boards as well. In fact, XMRC was forced to abandon explicitly gendered ads in March 2019.

### A.3.5 Other Internet Job Boards

58.com is China's largest online job board serving temporary and part time jobs. In contrast to the job sites discussed previously, employers on 58.com include a large number of individuals, not just firms. Most of the jobs posted have low skill requirements and are informal in nature (in the sense that they do not participate in the social insurance system). A search of 58.com, parallel to those of 51job, Zhaopin and Chinahr, indicated that both the words "man" and "woman" and their transformations are frequently used to request workers of a particular gender.[8] This may be due, in part, to

[8] Notably, this is despite the fact that 58's employer portal asks job posters, "Please do not include special symbols or any gender discriminating information".

workers' unwillingness to report individuals (as opposed to firms) for discrimination, and the small stakes involved in doing so. And again, demand for gendered ads may be higher due to the less-skilled nature of these jobs.

Finally, Yingjiesheng.com is a website that aggregates information about job openings for new university graduates from a number of sources, including the job boards described above. In addition to referring applicants to those job postings, Yingjiesheng provides information about the recruitment plans of firms attending campus job fairs, and about the recruitment plans posted by firms on their own websites. These plans frequently include explicit gender preferences, which can often vary within firms. For example, a firm's official, posted recruitment plan might say, "We are hiring 5 men for position A, 10 men for position B, and 5 women for position C".

This Appendix was prepared with the assistance of Naijia Wu, an undergraduate student at UCSB.

## A.4 Additional Tables and Figures

Table A.1: Descriptive Statistics: Ads in Full Sample

| | **Ad Requests Women** *F* **jobs** | **Gender not specified** *N* **jobs** | **Ad Requests Men** *M* **jobs** | **All Ads** |
|---|---|---|---|---|
| Education specified? | 0.946 | 0.886 | 0.931 | 0.906 |
| Education Requested (years), if specified | 12.83 | 12.74 | 11.71 | 12.57 |
| Tech School Requested? | 0.282 | 0.138 | 0.182 | 0.175 |
| Desired Age Range specified? | 0.576 | 0.321 | 0.530 | 0.408 |
| Desired Age, if Requested (midpoint of interval) | 26.37 | 29.54 | 30.32 | 28.85 |
| Experience Requested (years) | 0.837 | 1.158 | 1.348 | 1.129 |
| New Graduate Requested? | 0.036 | 0.017 | 0.019 | 0.021 |
| Wage Advertised? | 0.509 | 0.385 | 0.445 | 0.420 |
| Wage, if advertised (yuan/month, midpoint of | 2,013 | 2,730 | 2,515 | 2,520 |
| Number of positions specified? | 0.960 | 0.933 | 0.963 | 0.944 |
| Number of positions, if specified | 1.602 | 1.821 | 1.698 | 1.756 |
| Number of applicants | 58.99 | 42.45 | 36.96 | 44.96 |
| Sample Size | 8,324 (19.5%) | 26,769 (62.6%) | 7,651 (17.9%) | 42,744 (100%) |

Table A.2: Descriptive Statistics: Ads in Callback Sample

| | Ad Requests Women<br>*F* jobs | Gender not specified<br>*N* jobs | Ad Requests Men<br>*M* jobs | All Ads |
|---|---|---|---|---|
| Education specified? | 0.961 | 0.899 | 0.925 | 0.919 |
| Education Requested (years), if specified | 12.70 | 12.31 | 11.25 | 12.21 |
| Tech School Requested? | 0.301 | 0.165 | 0.206 | 0.207 |
| Desired Age Range specified? | 0.638 | 0.390 | 0.566 | 0.481 |
| Desired Age, if Requested (midpoint of interval) | 25.91 | 28.81 | 29.47 | 28.03 |
| Experience Requested (years) | 0.785 | 0.997 | 1.215 | 0.987 |
| New Graduate Requested? | 0.069 | 0.023 | 0.030 | 0.035 |
| Wage Advertised? | 0.638 | 0.557 | 0.556 | 0.576 |
| Wage, if advertised (yuan/month, midpoint of interval) | 2,001 | 2,658 | 2,439 | 2,446 |
| Number of positions specified? | 0.964 | 0.923 | 0.971 | 0.941 |
| Number of positions, if specified | 1.915 | 2.249 | 2.033 | 2.130 |
| Number of applicants | 79.49 | 62.56 | 46.55 | 63.66 |
| | | | | |
| Sample Size | 867 | 2,104 | 666 | 3,637 |

Table A.3: Descriptive Statistics: Applications in Callback Sample

| | Applications from | | All |
|---|---|---|---|
| | **Women** | **Men** | **Applications** |
| Education (years) | 14.56 | 14.11 | 14.35 |
| Completed Tech School? | 0.155 | 0.164 | 0.159 |
| Age (years) | 23.24 | 24.86 | 23.99 |
| Experience (years) | 2.674 | 3.886 | 3.230 |
| New Graduate? | 0.210 | 0.155 | 0.185 |
| Current wage listed? | 0.688 | 0.702 | 0.694 |
| Current wage, if listed (yuan/month) | 2,090 | 2,462 | 2,263 |
| Married (if marital status listed) | 0.140 | 0.215 | 0.174 |
| Occupational Qualification (*Zhicheng*)[1] | 1.086 | 1.403 | 1.231 |
| Myopic | 0.328 | 0.268 | 0.301 |
| Height (cm) | 160.6 | 171.5 | 165.6 |
| English CV available? | 0.145 | 0.104 | 0.126 |
| Number of Schools listed | 0.312 | 0.279 | 0.297 |
| Number of Experience Spells | 2.678 | 2.606 | 2.645 |
| Number of Certifications | 1.462 | 0.886 | 1.198 |
| Sample Size | 124,275 | 105,341 | 229,616 |

Notes:

Zhicheng is a nationally-recognized worker certification system that assigns an official rank (from one through six) to workers in almost every occupation. Ranks are based on education, experience and in some cases nationwide or province-wide exams.

Table A.4: Comparing XMRC Ads to the Employed, Private-Sector Population in Xiamen and Urban China

| Worker Characteristics | XMRC job ads | | Xiamen | Urban China |
|---|---|---|---|---|
| | Callback Sample | Full Sample | employed population | employed population |
| | (1) | (2) | (3) | (4) |
| Female (percent of gendered ads) | 56.56 | 52.11 | 46.75 | 44.23 |
| Education (years) | 12.21 | 12.57 | 10.56 | 10.59 |
| Age (years) | 28.03 | 28.85 | 30.77 | 32.64 |
| Monthly wage (RMB) | 2,446 | 2,520 | 2,185 | 2,147 |
| **Broad occupation (percent):** | | | | |
| Management | 1.68 | 1.99 | 4.3 | 4.59 |
| Sales and Procurement | 18.64 | 16.59 | 18.31 | 21.25 |
| Service Occupations | 15.40 | 12.29 | 21.68 | 22.28 |
| Professional/Technical | 27.30 | 29.92 | 7.99 | 8.21 |
| Production, Construction, Manufacturing | 29.39 | 31.37 | 47.71 | 43.68 |
| Other | 7.59 | 7.83 | -- | -- |
| | | | | |
| Number of observations | 3,637 | 42,744 | 1,163 | 99,768 |

Notes:

1. Employment data are from the 2005 Census, 1% sample, persons currently living in urban regions, who are currently employed in the private sector (i.e. excluding SOEs, government and collectives).

2. "Urban China" comprises the four municipalities directly under the jurisdiction of the central government (Beijing, Shanghai, Tianjin and Chongqing) plus the 15 sub-provincial cities: Changchun, Chengdu, Dalian, Guangzhou, Hangzhou, Harbin, Jinan, Nanjing, Ningbo, Qingdao, Shenyang, Shenzhen, Wuhan, Xiamen, and Xi'an.

3. Chinese wages have been adjusted for per capita GDP growth between 2005 and 2010 using IMF GDP statistics.

4. The occupational classification system used by XMRC uses 37 categories that were created by the website; mapping these into Census categories is a fairly subjective exercise. Therefore we only attempt this at a very high level of aggregation, and we recommend treating the occupational statistics with caution.

Table A.5: Matching, Compliance and Enforcement Rates for Other Forms of Mismatch

| | Matching | Compliance | Enforcement |
|---|---|---|---|
| | Share of callbacks that match the employer's request | Share of applications that match the employer's request | Share of mismatched applications that are rejected |
| | (1) | (2) | (3) |
| **A. Matching on Five Dimensions** | | | |
| 1. Gender | 0.948 | 0.925 | 0.947 |
| 2. Education | 0.436 | 0.444 | 0.917 |
| 3. Experience | 0.602 | 0.597 | 0.917 |
| 4. Wage | 0.495 | 0.501 | 0.916 |
| 5. Age (within requested range) | 0.748 | 0.734 | 0.925 |
| **B. Alternative Measures of Age Matching** | | | |
| 6. Age (within range plus one year) | 0.839 | 0.830 | 0.925 |
| 7. Age (within range plus 2 years) | 0.908 | 0.899 | 0.928 |
| 8. Age (within 2 years of the median) | 0.317 | 0.321 | 0.921 |
| 9. Age (within 3 years of the median) | 0.442 | 0.449 | 0.920 |
| 10. Age (within 4 years of the median) | 0.562 | 0.565 | 0.920 |

Notes:

1. Row 2: Education matching means the candidate's education falls into the education category that is requested in the ad. The five education categories are: primary or less (6 years), junior middle school (9 years), high school (12 years), college or technical school (15 years) and university (16 years).
2. Row 3: Experience matching means the candidate's experience equals the amount requested in the ad, or exceeds the request by no more than three years.
3. Row 4: Wage matching means the applicant's current wage is in the same wage category as the job's advertised wage. The wage categories (in RMB/month) are "around 1000", 1000-1999, 2000-2999, 3000-3999, 4000-4999, 5000-5999, 6000-7999, and 8000-9999. Since 99 percent of offered and current wages are below 6000, this means that the candidate's wage is, on average, within about 1000 RMB/month of the offered wage, or within about one standard deviation.
4. Row 5: A worker is age-matched if her age is within the employer's requested age range.
5. Row 6: A worker is age-matched if her age is within the employer's requested age range, or is one year outside the requested range.

6. Row 7: a worker is age-matched if her age is within the employer's requested age range, or is up to two years outside the requested range.

7. Rows 8–10: A worker is age-matched if her age is within 2, 3 or 4 years of the midpoint of the employer's requested age range.

Table A.6: Effects of Gender Requests on the Share of Female Applications Received ($\alpha$) Callback Sample Only

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Ad requests men ($M$) | -0.3680*** | -0.3270*** | -0.2350*** | -0.1368*** | -0.1034*** |
| | (0.020) | (0.019) | (0.017) | (0.019) | (0.032) |
| Ad requests women ($F$) | 0.4790*** | 0.4243*** | 0.3603*** | 0.2037*** | 0.2401*** |
| | (0.014) | (0.016) | (0.016) | (0.012) | (0.024) |
| Primary School | | 0.0292 | -0.0113 | 0.0063 | 0.0502 |
| | | (0.034) | (0.029) | (0.019) | (0.032) |
| Middle School | | -0.0683* | -0.0518** | -0.0087 | 0.0346 |
| | | (0.036) | (0.023) | (0.021) | (0.027) |
| Tech School | | 0.0587** | 0.0287 | -0.0125 | -0.0322 |
| | | (0.026) | (0.021) | (0.016) | (0.027) |
| Post-secondary | | 0.1275*** | 0.0600*** | 0.0033 | 0.0215 |
| | | (0.024) | (0.020) | (0.016) | (0.029) |
| University | | 0.1062*** | 0.0361 | -0.0113 | -0.0757 |
| | | (0.038) | (0.027) | (0.025) | (0.064) |
| Number of positions advertised | | -1.2386*** | -0.8558*** | 0.3736 | 0.5534 |
| | | (0.330) | (0.268) | (0.288) | (0.512) |
| Occupation Fixed Effects | | | Y | Y | Y |
| Job Title Fixed Effects | | | | Y | Y |
| Firm Fixed Effects | | | | | Y |
| $N$ (ads) | 3,637 | 3,637 | 3,637 | 3,637 | 3,637 |
| 'Effective' $N$ | 3,637 | 3,637 | 3,637 | 1,627 | 840 |
| $R^2$ | 0.571 | 0.620 | 0.738 | 0.936 | 0.980 |

Notes:

1. In addition to the covariates shown, columns 2-5 also control for the following job ad characteristics: requested experience level (quadratic), requested age level (quadratic in midpoint of range), advertised wage (quadratic in midpoint of bin; 8 bins), dummy for whether new graduate requested, number of positions advertised, plus dummies for missing education, age, wage and number of positions.

2. All regressions are weighted by the total number of applications received.

3. 'Effective' N excludes job titles, and firm IDs that only appear in one ad in columns 4, and 5 respectively.

4. Table A.6 replicates Table 1.3 for the sample of job ads for which we observe callback information. The most saturated specification we can estimate in this smaller sample replicates column 5, where firm and job title fixed effects are entered separately. The

estimated effects of male and female labels of -.103 and .240 are very similar to Table 1.3's estimates of -.120 and .234; all of these coefficients are highly statistically significant. Standard errors in parentheses, clustered by firm. *** p<0.01, ** p<0.05, * p<0.1

Table A.7

Table A.7a: Descriptive Statistics: 1448 ads in Table 1.3 Column 6

| | **Ad Requests Women** <br> ***F* jobs** | **Gender not specified** <br> ***N* jobs** | **Ad Requests Men** <br> ***M* jobs** | **All Ads** |
|---|---|---|---|---|
| Education specified? | 0.947 | 0.880 | 0.906 | 0.900 |
| Education Requested (years), if specified | 12.81 | 12.44 | 11.40 | 12.32 |
| Tech School Requested? | 0.291 | 0.153 | 0.243 | 0.203 |
| Desired Age Range specified? | 0.591 | 0.333 | 0.483 | 0.420 |
| Desired Age, if Requested (midpoint of interval) | 25.74 | 28.51 | 29.55 | 27.87 |
| Experience Requested (years) | 0.749 | 0.967 | 1.259 | 0.976 |
| New Graduate Requested? | 0.068 | 0.018 | 0.017 | 0.029 |
| Wage Advertised? | 0.495 | 0.398 | 0.500 | 0.440 |
| Wage, if advertised (yuan/month, midpoint of | 1,841 | 2,379 | 2,357 | 2,239 |
| Number of positions specified? | 0.960 | 0.921 | 0.979 | 0.941 |
| Number of positions, if specified | 1.430 | 1.801 | 1.657 | 1.690 |
| Number of applicants | 43.99 | 38.67 | 28.63 | 37.87 |
| Sample Size | 323 (22.3%) | 839 (57.9%) | 286 (19.8%) | 1,448 (100%) |

Notes:

Compared to the full sample of 42,744 job ads (Table A.1), the 1,448 job ads that identify our preferred estimates of compliance (column 6 of Table 1.3) are:

• slightly more likely to request men and women (1.8 and 2.8 percentage points respectively)

• request slightly less education (0.25 years)

• -request slightly younger workers (about one year)

• -request a little less experience (0.15 years)

• -offer 11.2% lower wages (2520-2239)/2520.

Table A.7b: The five most common occupations in the 'identifying' and full samples

| | "Identifying" Sample ($N$=1,448 ads) | | Full Sample ($N$= 42,744 ads) | |
|---|---|---|---|---|
| Rank | Occupation Name | Share | Occupation Name | Share |
| 1 | Construction | 11.74% | Sales | 11.12% |
| 2 | Sales | 10.08% | Construction | 10.44% |
| 3 | Administration | 9.67% | Software | 8.45% |
| 4 | Clothes | 7.18% | Manufacture | 7.83% |
| 5 | Manufacture | 6.84% | Administration | 6.94% |

Notes:

Four of the five most common occupational categories applied to in the identifying sample are in the top five in the overall sample: Construction, Sales, Manufacturing and Administration. These four categories comprise 38 percent of the identifying sample and 36 percent of the overall sample.

Table A.7c: The ten most common job titles in 'identifying' and full samples

| | "Identifying" Sample ($N$=1,448 ads) | | Full Sample ($N$= 42,744 ads) | |
|---|---|---|---|---|
| Rank | Job Title Name | Share | Job Title Name | Share |
| 1 | warehouse management | 3.87% | warehouse management | 1.39% |
| 2 | international trade business person | 3.52% | accountant | 1.29% |
| 3 | driver | 3.04% | driver | 1.13% |
| 4 | business person | 3.04% | business person | 0.99% |
| 5 | clerk | 2.90% | cashier/teller | 0.97% |
| 6 | accountant | 2.62% | clerk | 0.90% |
| 7 | cashier/teller | 1.38% | international trade business person | 0.77% |
| 8 | business assistant | 1.24% | general manager assistant | 0.65% |
| 9 | general manager assistant | 1.17% | business assistant | 0.57% |
| 10 | technician | 1.17% | front desk clerk | 0.47% |

Notes:

Nine of the ten most common job titles in the identifying samples are in the top ten job titles overall. Together these titles account for 23.9 percent of the identifying job titles, and 9.1 percent of all job titles.

Thus (as we would expect), identification comes disproportionately from the most common job titles, since the chances of observing different gender requests attached to the same title inside the same firm increase with the prevalence of the title.

Table A.8

Table A.8a: Men's Selection into Gender-Mismatched Applications

| | Destination of the Application | | |
|---|---|---|---|
| | F job (gender-mismatched) | N and M jobs (gender-matched) | Difference (1) – (2) |
| | (1) | (2) | (3) |
| **1. Share of applications satisfying job's posted requirements:** | | | |
| Matches Education requirement: | 0.440 | 0.466 | -0.026*** |
| Education less than requested | 0.073 | 0.066 | 0.008** |
| Education more than requested | 0.486 | 0.468 | 0.018** |
| Meets or Exceeds Experience requirement | 0.889 | 0.874 | 0.015*** |
| Matches Age requirement: | 0.832 | 0.855 | -0.023*** |
| Age less than requested | 0.132 | 0.126 | 0.006 |
| Age more than requested | 0.037 | 0.020 | 0.017*** |
| | | | |
| **2. Mean callback rates in the:** | | | |
| Occupation applied to | 0.081 | 0.090 | -0.009*** |
| Job title applied to | 0.081 | 0.095 | -0.015*** |
| | | | |
| **3. Total number of:** | | | |
| Applications submitted, per applicant | 5.703 | 2.497 | 3.206*** |
| Occupations applied to, per applicant | 3.126 | 1.768 | 1.358*** |
| Job Titles applied to, per applicant | 4.831 | 2.288 | 2.543*** |

Notes:

1. Men who apply to jobs requesting women are less likely to match the job's education and age requirements, and more likely to meet the job's experience requirement than men applying to gender-matched jobs.

2. Men who apply to jobs requesting women send their (entire) application packet to occupations and job titles with lower callback rates, compared to men who apply to gender-matched jobs.

3. Men who apply to gender-mismatched jobs submit many more applications overall, and apply to a much greater variety of occupations and job titles than men who apply to gender-matched jobs.

4. *** p<0.01, ** p<0.05, * p<0.1, based on two-sample t-test with equal variance. (No significance levels change if we use unequal variances).

Table A.8b: Women's Selection into Gender-Mismatched Applications

| | Destination of the Application | | |
|---|---|---|---|
| | M job (gender-mismatched) | N and F jobs (gender-matched) | Difference (1) – (2) |
| | (1) | (2) | (3) |
| **1. Share of applications satisfying job's posted requirements:** | | | |
| Matches Education requirement: | 0.455 | 0.481 | -0.026*** |
| Education less than requested | 0.102 | 0.077 | 0.025*** |
| Education more than requested | 0.443 | 0.442 | 0.001 |
| Meets or Exceeds Experience requirement | 0.886 | 0.844 | 0.042*** |
| Matches Age requirement: | 0.795 | 0.826 | -0.031*** |
| Age less than requested | 0.177 | 0.149 | 0.028*** |
| Age more than requested | 0.028 | 0.025 | 0.003 |
| | | | |
| **2. Mean callback rates in the:** | | | |
| Occupation applied to | 0.082 | 0.079 | 0.003*** |
| Job title applied to | 0.052 | 0.075 | -0.023*** |
| | | | |
| **3. Total number of:** | | | |
| Applications submitted, per applicant | 6.743 | 3.377 | 3.366*** |
| Occupations applied to, per applicant | 3.374 | 2.118 | 1.256*** |
| Job Titles applied to, per applicant | 5.637 | 2.950 | 2.687*** |

Notes:

1. Women who apply to jobs requesting men are less likely to match the job's education and age requirements, and more likely to meet the job's experience requirement than women applying to gender-matched jobs. All these differences are quite small, however.
2. Women who apply to jobs requesting men send their (entire) application packet to job titles with lower callback rates, but occupations with slightly higher callback rates, compared to women who apply to gender-matched jobs.
3. Women who apply to gender-mismatched jobs submit many more applications overall, and apply to a much greater variety of occupations and job titles than women who apply to gender-matched jobs.

4. *** $p<0.01$, ** $p<0.05$, * $p<0.1$, based on two-sample t-test with equal variance. (No significance levels change if we use unequal variances).

Table A.9: Effects of Gender Requests on Skill Requirements and Offered Wages

| | (1) Education Requested | (2) Experience Requested | (3) Offered Wage |
|---|---|---|---|
| Ad requests men (*M*) | 0.0358 | 0.0465 | 0.0014 |
| | (0.130) | (0.116) | (0.066) |
| Ad requests women (*F*) | -0.1011 | -0.0872 | -0.1921*** |
| | (0.109) | (0.079) | (0.060) |
| Age | 0.3270** | -0.0475 | -0.0572 |
| | (0.138) | (0.136) | (0.093) |
| Age^2 | -0.0487** | 0.0317 | 0.0174 |
| | (0.023) | (0.023) | (0.016) |
| Primary School | | | 0.1572 |
| | | | (0.117) |
| Middle School | | | -0.0793 |
| | | | (0.099) |
| Tech School | | | 0.0074 |
| | | | (0.060) |
| Post-secondary | | | 0.0783 |
| | | | (0.072) |
| University | | | 0.4333** |
| | | | (0.175) |
| Exp | | | 0.0117 |
| | | | (0.031) |
| Exp^2 | | | 0.0507 |
| | | | (0.045) |
| Occupation Fixed Effects | Y | Y | Y |
| Job Title Fixed Effects | Y | Y | Y |
| Firm Fixed Effects | Y | Y | Y |
| Mean of Y | 12.85 | 2.43 | 2.46 |
| N (ads) | 16,724 | 10,395 | 8,852 |
| "Effective" *N* | 7,666 | 3,957 | 3,234 |
| $R^2$ | 0.765 | 0.827 | 0.874 |

Standard errors in parentheses, clustered by firm. *** p<0.01, ** p<0.05, * p<0.1

Notes:

1. Table A.9 replicates column 5 in Table 1.3, using skill requirements and the offered wage as outcome variables, and removing the controls for posted wages, number of positions advertised and requests for new graduates. Ads with a missing age requirement are excluded from all regressions.

2. In column 1, the dependent variable is requested years of education; thus ads with missing education requirements are dropped from the sample. Controls for requested education and experience are also excluded.

3. In column 2, the dependent variable is requested years of experience; thus ads with no explicit experience requirement are dropped from the sample. Controls for requested education and experience are also excluded.

4. In column 3, the dependent variable is offered wage in thousands of yuan per month; thus ads that do not post a wage are dropped from the sample.

Once we control for detailed duties (with job title fixed effects) jobs that explicitly request men or women do not require higher or lower levels of skill –as measured by education or experience– than non-gendered job ads. Even within job titles, however, ads explicitly requesting women pay (.1921/2.46=) 7.8 percent less, a difference which is highly statistically significant.

Figure A.1: Distribution of Leave-Out-One-Title Estimates of Gender Request Effects on Female Applicant Shares

Figure A.1a: Effect of a Request for Men

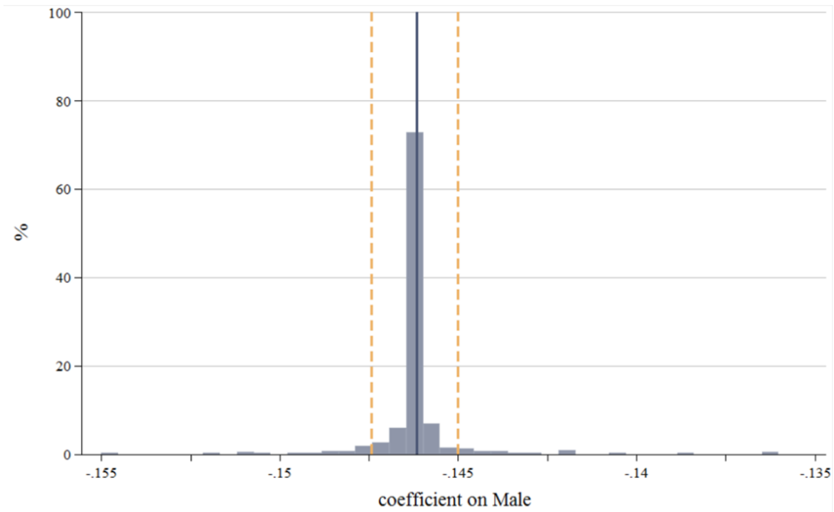
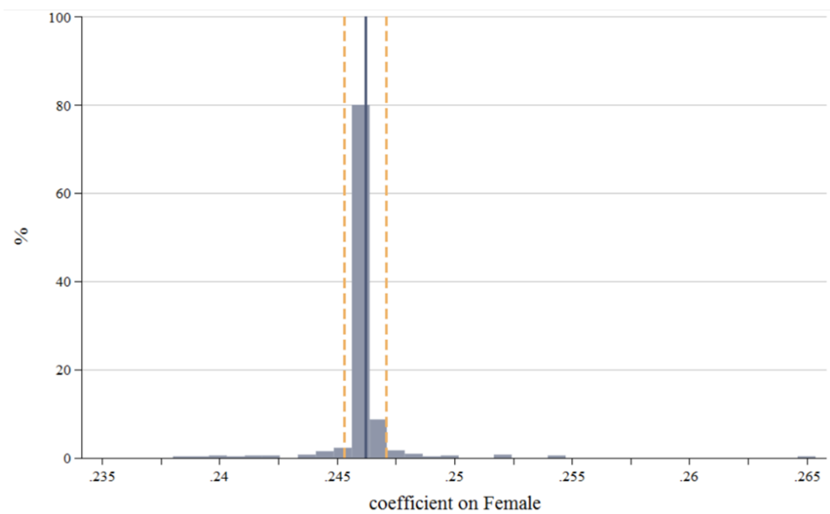



Figure A.1b: Effect of a Request for Women

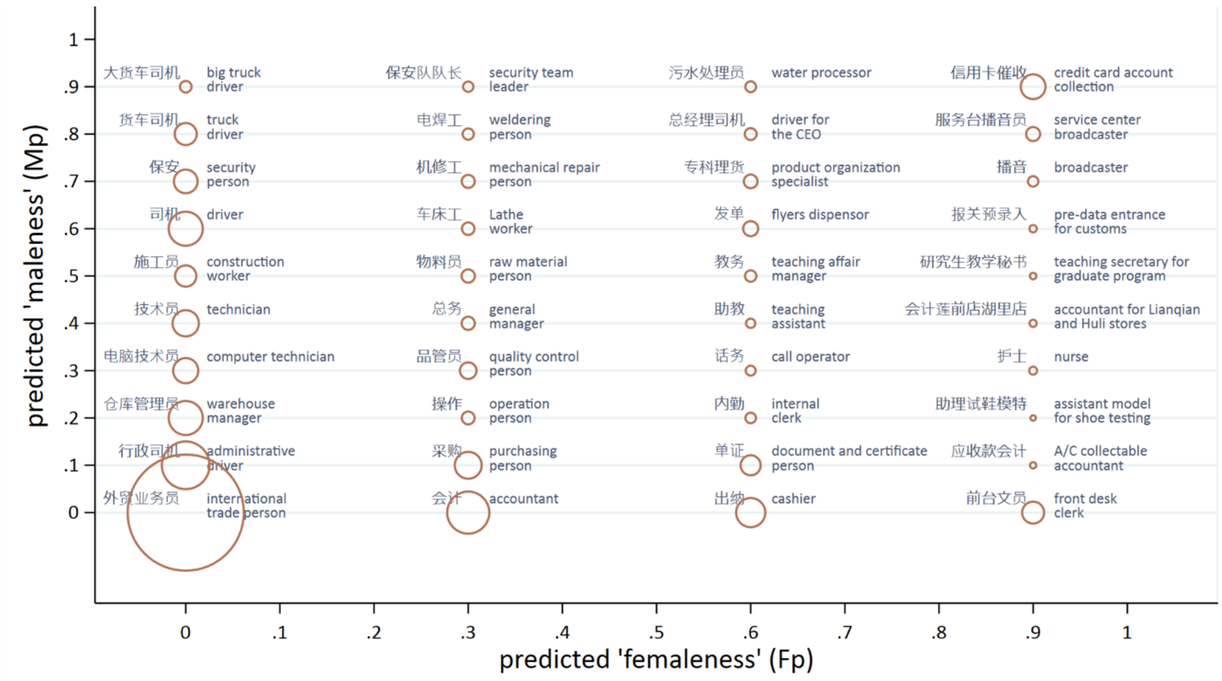



Notes:

1. There two figures present estimates of the "Ad requests men" and "Ad requests women" coefficients in column 6 of Table 1.3.
2. These coefficients are identified by 416 distinct job titles; the Figures report the distribution of estimates when one job title is dropped at a time.
3. Vertical solid line represents the entire-sample estimate; vertical dashed lines show the 5th and 95 percentiles of the estimates.

4. All estimates of the request-male effect are between -.155 and -.136 and statistically significant (p<.01). All estimates of the request-female effect are between .238 and .265 and statistically significant (p<.01).

Figure A.2: Selected Job Titles, by Predicted 'Maleness' ($M_p$) and 'Femaleness' ($F_p$)




Notes:

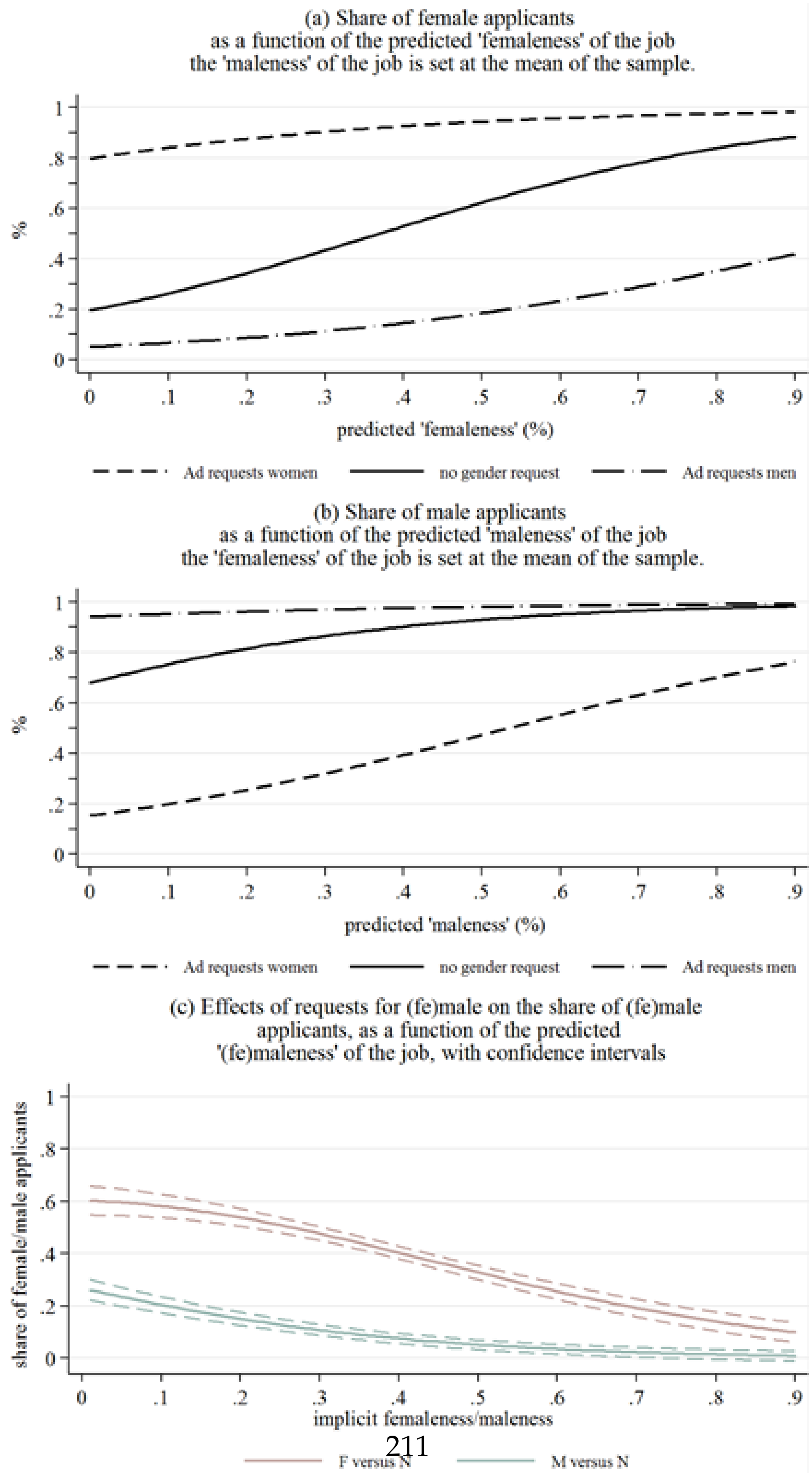
1. Symbol size is proportional to the number of unique job ads.

2. The job titles shown are the job titles that correspond to the largest number of applications in each cell.

3. The forty cells in the figure are defined by four predicted 'femaleness' ranges ([0,0.1],[.3,.4],[.6,.7],[.9,1]) and ten predicted 'maleness' ranges ([0,0.1],[.1,.2],,[.9,1]).

Figure A.2 shows that "front desk clerks" and "big truck drivers" are typically female and male jobs, respectively. Other jobs, like "credit card account collection" express gender preferences frequently, but prefer females in some postings and males in others. Finally, jobs like "international trade person" rarely express an explicit gender preference; thus, the predicted 'maleness' and 'femaleness' of these jobs are both very low.

Figure A.3: Effects of Gender Requests and Predicted Genderness of the Job Ads on the Gender Mix of Applications Received (Full Ad Sample, log-odds specification)



Notes:

1. Figure A.3 shows predicted values of the female share of applicants ($\alpha$) from a specification identical to Figure 1, with the following changes:

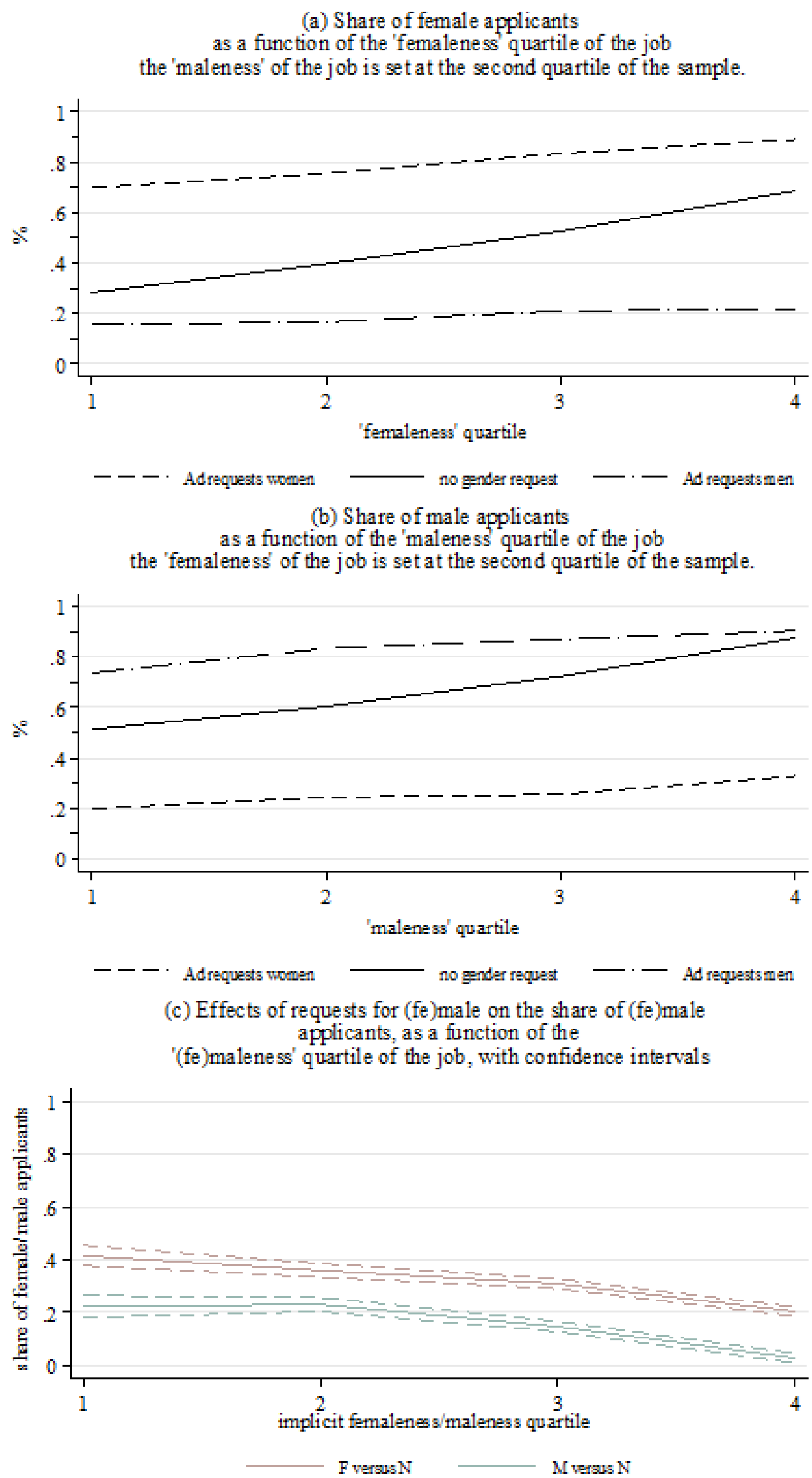
   (1) The dependent variable $\alpha$, is now log $\alpha/(1-\alpha)$. 'Corner' values of $\alpha$ are accommodated by setting $\alpha = 0.5/A$ when $\alpha = 0$ and setting $\alpha = (A - 0.5)/A$ when $\alpha = 1$, where A is the total number of applications to the ad.

   (2) the quartics in $F_p$ and $M_p$ (each interacted with F, N and M) are replaced by linear terms (again interacted with F, N and M ).

2. As in Figure 1, predictions in part (a) hold $M_p$ at its mean, and predictions in part (b) hold $F_p$ at its mean. All other characteristics are set at their means.

3. The regression is weighted by the number of applications to each ad, and standard errors are clustered at the firm level.

Figure A.4: Effects of Gender Requests and Predicted Genderness of the Job Ads on the Gender Mix of Applications Received (Full Ad Sample, quartile dummy specification)



Notes:

1. Figure A.4 replaces the quartic in Figure 1 by a set of fixed effects for quartiles of the predicted maleness ($M_p$) and femaleness ($F_p$).

2. Quartiles of $M_p$ are: 0.0137, 0.0856, 0.3948; Quartiles of $F_p$ are: 0.0308, 0.1462, 0.7235.

3. Predictions in part (a), which shows the effect of implicit femaleness ($F_p$) quartiles, hold ($M_p$) at the second quartile. Predictions in part (b), which depicts the implicit maleness ($M_p$) quartiles, hold $F_p$ at the second quartile. All other characteristics are set at their means. The regression is weighted by the number of applications to each ad, and standard errors are clustered at the firm level.

## A.5 Effects of Explicit Gender Requests on the Number and Quality of Applications

Table A.10 replicates Table 1.3, using the total number of applications received as the outcome variable. Tables A.11 - A.16 do the same for a variety of measures of the average quality of the applicant pool.

In some of the uncontrolled regressions of column 1, a number of effects are estimated, which confirm known features of the data: there are more female jobseekers than men on the board; gendered job ads are more common in unskilled positions, and men tend to have more experience than women.

Once job titles are controlled for, however (columns 5 and 6),

-there is some evidence that employers pay a price in applicant numbers when they advertise a gender preference (though the estimates are imprecise in column 6)

-there is no detectable effect of gender requests on mean applicant education and experience (Tables A.11 and A.12)

-all the estimated effects of gender requests on match quality (Tables A.13 - A.16) are small and statistically insignificant

Table A.10: Effects of Gender Requests on the Number of Applications Received

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Ad requests men ($M$) | -5.2452*** | -2.5686*** | -1.6847** | -5.6321*** | -6.1424*** | -12.3542** |
| | (0.859) | (0.838) | (0.821) | (1.261) | (1.487) | (5.789) |
| Ad requests women ($F$) | 16.9273*** | 10.6785*** | 0.9783 | -11.7076*** | -11.6047*** | -13.5509* |
| | (1.114) | (1.158) | (1.147) | (1.890) | (1.909) | (7.660) |
| Primary School | | -10.3082*** | -10.8216*** | -10.6052*** | -11.0666*** | 10.2855 |
| | | (1.216) | (1.184) | (1.902) | (2.004) | (8.900) |
| Middle School | | -12.6842*** | -10.8267*** | -3.8782** | -3.8040* | -0.5427 |
| | | (1.349) | (1.290) | (1.956) | (2.186) | (8.779) |
| Tech School | | 9.8202*** | 7.7057*** | 7.7421*** | 7.2772*** | 20.4669* |
| | | (1.301) | (1.250) | (1.796) | (1.961) | (9.052) |
| Post-secondary | | 16.8829*** | 12.6597*** | 13.2305*** | 8.8733*** | 9.5534 |
| | | (1.234) | (1.192) | (1.872) | (2.000) | (9.722) |
| University | | 11.6652*** | 5.5642*** | 13.7460*** | 4.0007 | 9.9381 |
| | | (1.617) | (1.565) | (2.620) | (2.762) | (10.497) |
| Number of positions advertised | | -70.5760*** | 76.4291*** | 232.0516*** | 252.2924*** | 112.1543 |
| | | (20.774) | (20.559) | (31.588) | (38.571) | (150.076) |
| Occupation Fixed Effects | | | Y | Y | Y | Y |
| Job Title Fixed Effects | | | | Y | Y | |
| Firm Fixed Effects | | | | | Y | |
| Title*Firm Fixed Effects | | | | | | Y |
| "Effective" $N$ | 42,744 | 42,744 | 42,744 | 25,438 | 23,819 | 1,448 |
| $R^2$ | 0.012 | 0.044 | 0.117 | 0.281 | 0.546 | 0.675 |

Standard errors in parentheses, clustered by firm. *** p<0.01, ** p<0.05, * p<0.1

Notes:

1. Dependent Variable: the total number of applications received, mean = 44.14.
2. Relative to the mean of 44.14 applications, column 6 indicates that adding a request for men (women) to a job ad reduces the number of applications received by 28 (31) percent.

Table A.11: Effects of Employers' Gender Requests on the Mean Education of Applicants

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Ad requests men ($M$) | -1.0935*** | -0.5825*** | -0.4337*** | 0.0110 | 0.0135 | -0.0344 |
| | (0.035) | (0.022) | (0.021) | (0.011) | (0.012) | (0.076) |
| Ad requests women ($F$) | -0.1752*** | -0.0568*** | -0.0876*** | -0.1138*** | -0.1100*** | -0.0815 |
| | (0.018) | (0.013) | (0.013) | (0.012) | (0.011) | (0.071) |
| Primary School | | 0.2527*** | 0.2087*** | 0.0902*** | 0.0898*** | -0.0989 |
| | | (0.043) | (0.038) | (0.024) | (0.020) | (0.103) |
| Middle School | | -0.8070*** | -0.7150*** | -0.1319*** | -0.1111*** | -0.2152** |
| | | (0.044) | (0.040) | (0.019) | (0.020) | (0.106) |
| Tech School | | 0.5596*** | 0.4658*** | 0.0519*** | 0.0384** | -0.0256 |
| | | (0.028) | (0.026) | (0.015) | (0.016) | (0.100) |
| Post-secondary | | 1.5631*** | 1.3783*** | 0.6044*** | 0.5571*** | 0.5326*** |
| | | (0.024) | (0.023) | (0.015) | (0.017) | (0.116) |
| University | | 2.0975*** | 1.8389*** | 0.9070*** | 0.8076*** | 0.6571*** |
| | | (0.026) | (0.026) | (0.018) | (0.021) | (0.136) |
| Number of positions advertised | | 1.2563*** | 0.4973 | -0.0202 | -0.5632* | -0.6903 |
| | | (0.380) | (0.347) | (0.302) | (0.322) | (1.709) |
| Occupation Fixed Effects | | | Y | Y | Y | Y |
| Job Title Fixed Effects | | | | Y | Y | |
| Firm Fixed Effects | | | | | Y | |
| Title*Firm Fixed Effects | | | | | | Y |
| "Effective" $N$ | 42,744 | 42,744 | 42,744 | 25,438 | 23,819 | 1,448 |
| $R^2$ | 0.119 | 0.594 | 0.650 | 0.916 | 0.947 | 0.971 |

Standard errors in parentheses, clustered by firm. *** p<0.01, ** p<0.05, * p<0.1

Notes:

1. Dependent Variable: the mean education of all the applicants, mean = 14.21.

2. Regressions are weighted by the number of applications to the ad.

Table A.12: Effects of Gender Requests on the Mean Experience of Applicants

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Ad requests men ($M$) | 1.4249*** | 0.7556*** | 0.6249*** | 0.0886*** | 0.0975*** | 0.3780 |
| | (0.061) | (0.046) | (0.046) | (0.031) | (0.035) | (0.288) |
| Ad requests women ($F$) | -0.7793*** | -0.3207*** | -0.3613*** | -0.0284 | -0.0659** | 0.1091 |
| | (0.037) | (0.030) | (0.030) | (0.026) | (0.032) | (0.120) |
| Primary School | | 0.0071 | 0.0458 | -0.0071 | -0.0170 | 0.0311 |
| | | (0.067) | (0.061) | (0.040) | (0.046) | (0.220) |
| Middle School | | 0.5254*** | 0.3780*** | -0.1189** | -0.1241** | 0.6836** |
| | | (0.078) | (0.077) | (0.053) | (0.058) | (0.299) |
| Tech School | | -0.7392*** | -0.6734*** | -0.1489*** | -0.1578*** | -0.1611 |
| | | (0.049) | (0.048) | (0.030) | (0.038) | (0.171) |
| Post-secondary | | -1.3039*** | -1.1177*** | -0.4632*** | -0.4501*** | -0.5061*** |
| | | (0.045) | (0.044) | (0.032) | (0.040) | (0.176) |
| University | | -1.6051*** | -1.3712*** | -0.6387*** | -0.6401*** | -0.6166** |
| | | (0.064) | (0.064) | (0.048) | (0.057) | (0.289) |
| Number of positions advertised | | -20.3136*** | -16.3763*** | -5.0431*** | -3.3648*** | -2.3730 |
| | | (0.834) | (0.771) | (0.674) | (0.750) | (3.470) |
| Occupation Fixed Effects | | | Y | Y | Y | Y |
| Job Title Fixed Effects | | | | Y | Y | |
| Firm Fixed Effects | | | | | Y | |
| Title*Firm Fixed Effects | | | | | | Y |
| "Effective" $N$ | 42,744 | 42,744 | 42,744 | 25,438 | 23,819 | 1,448 |
| $R^2$ | 0.090 | 0.419 | 0.478 | 0.855 | 0.905 | 0.941 |

Standard errors in parentheses, clustered by firm. *** p<0.01, ** p<0.05, * p<0.1

Notes:

1. Dependent Variable: the mean experience of all the applicants, mean = 4.13.
2. Regressions are weighted by the number of applications to the ad.

Table A.13: Effects of Gender Requests on the Share of Applicants Satisfying the Job's Education Requirement

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Ad requests men ($M$) | 0.0036 | -0.0274*** | -0.0199*** | -0.0033 | -0.0020 | 0.0160 |
| | (0.003) | (0.002) | (0.002) | (0.003) | (0.003) | (0.013) |
| Ad requests women ($F$) | 0.0083** | -0.0130*** | -0.0161*** | -0.0111*** | -0.0120*** | 0.0011 |
| | (0.003) | (0.002) | (0.002) | (0.002) | (0.003) | (0.009) |
| Primary School | | 0.0275*** | 0.0252*** | 0.0230*** | 0.0264*** | 0.0192 |
| | | (0.002) | (0.002) | (0.003) | (0.003) | (0.013) |
| Middle School | | 0.0406*** | 0.0460*** | 0.0606*** | 0.0655*** | 0.0437*** |
| | | (0.002) | (0.002) | (0.003) | (0.004) | (0.016) |
| Tech School | | 0.0236*** | 0.0195*** | 0.0118*** | 0.0129*** | 0.0207* |
| | | (0.002) | (0.002) | (0.002) | (0.003) | (0.011) |
| Post-secondary | | -0.0739*** | -0.0853*** | -0.1200*** | -0.1217*** | -0.1100*** |
| | | (0.002) | (0.002) | (0.003) | (0.004) | (0.015) |
| University | | -0.3047*** | -0.3244*** | -0.4123*** | -0.4140*** | -0.3193*** |
| | | (0.005) | (0.005) | (0.006) | (0.006) | (0.042) |
| Number of positions advertised | | 0.1947*** | 0.1221*** | -0.0458 | -0.0378 | 0.4149 |
| | | (0.044) | (0.038) | (0.055) | (0.073) | (0.327) |
| Occupation Fixed Effects | | | Y | Y | Y | Y |
| Job Title Fixed Effects | | | | Y | Y | |
| Firm Fixed Effects | | | | | Y | |
| Title*Firm Fixed Effects | | | | | | Y |
| "Effective" $N$ | 42,744 | 42,744 | 42,744 | 25,438 | 23,819 | 1,448 |
| $R^2$ | 0.001 | 0.567 | 0.594 | 0.755 | 0.838 | 0.895 |

Standard errors in parentheses, clustered by firm. *** p<0.01, ** p<0.05, * p<0.1

Notes:

1. Dependent Variable: the share of applicants that satisfy the job's education requirement, mean = .9178.

2. Regressions are weighted by the number of applications to the ad.

Table A.14: Effects of Gender Requests on the Share of Applicants Satisfying the Job's Experience Requirement

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Ad requests men ($M$) | 0.0292*** | 0.0317*** | 0.0278*** | 0.0155*** | 0.0158*** | 0.0263 |
| | (0.004) | (0.003) | (0.003) | (0.004) | (0.004) | (0.039) |
| Ad requests women ($F$) | -0.0077* | -0.0135*** | -0.0145*** | -0.0016 | -0.0006 | 0.0015 |
| | (0.004) | (0.003) | (0.003) | (0.003) | (0.004) | (0.017) |
| Primary School | | -0.0197*** | -0.0169*** | -0.0142*** | -0.0154*** | -0.0142 |
| | | (0.003) | (0.003) | (0.005) | (0.005) | (0.020) |
| Middle School | | 0.0183*** | 0.0143*** | 0.0033 | -0.0004 | -0.0074 |
| | | (0.005) | (0.005) | (0.007) | (0.007) | (0.021) |
| Tech School | | -0.0395*** | -0.0365*** | -0.0185*** | -0.0191*** | -0.0339* |
| | | (0.004) | (0.004) | (0.004) | (0.004) | (0.019) |
| Post-secondary | | -0.0685*** | -0.0598*** | -0.0347*** | -0.0389*** | -0.0564*** |
| | | (0.003) | (0.003) | (0.004) | (0.004) | (0.019) |
| University | | -0.0797*** | -0.0690*** | -0.0418*** | -0.0495*** | -0.0644 |
| | | (0.005) | (0.005) | (0.006) | (0.007) | (0.042) |
| Number of positions advertised | | -0.5309*** | -0.3733*** | -0.0715 | -0.1033 | 0.4454 |
| | | (0.066) | (0.065) | (0.087) | (0.100) | (0.574) |
| Occupation Fixed Effects | | | Y | Y | Y | Y |
| Job Title Fixed Effects | | | | Y | Y | |
| Firm Fixed Effects | | | | | Y | |
| Title*Firm Fixed Effects | | | | | | Y |
| "Effective" $N$ | 42,744 | 42,744 | 42,744 | 25,438 | 23,819 | 1,448 |
| $R^2$ | 0.004 | 0.590 | 0.606 | 0.748 | 0.836 | 0.908 |

Standard errors in parentheses, clustered by firm. *** p<0.01, ** p<0.05, * p<0.1

Notes:

1. Dependent Variable: the share of applicants that satisfy the job's experience requirement, mean = .8741.

2. Regressions are weighted by the number of applications to the ad.

Table A.15: Effects of Gender Requests on the Share of Applicants Satisfying the Job's Age Requirement

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Ad requests men ($M$) | -0.0578*** | 0.0002 | -0.0053 | -0.0082 | -0.0094 | -0.0484 |
| | (0.006) | (0.004) | (0.004) | (0.006) | (0.007) | (0.043) |
| Ad requests women ($F$) | -0.0393*** | -0.0240*** | -0.0189*** | -0.0070 | -0.0017 | -0.0300 |
| | (0.006) | (0.004) | (0.004) | (0.006) | (0.007) | (0.028) |
| Primary School | | 0.0079 | 0.0097 | 0.0047 | 0.0169* | 0.0023 |
| | | (0.006) | (0.006) | (0.008) | (0.009) | (0.051) |
| Middle School | | 0.0313*** | 0.0307*** | 0.0067 | 0.0132 | -0.0909 |
| | | (0.007) | (0.007) | (0.010) | (0.012) | (0.084) |
| Tech School | | -0.0239*** | -0.0221*** | -0.0184** | -0.0095 | -0.0208 |
| | | (0.006) | (0.006) | (0.009) | (0.008) | (0.048) |
| Post-secondary | | -0.0346*** | -0.0305*** | -0.0256*** | -0.0152* | -0.0249 |
| | | (0.005) | (0.005) | (0.008) | (0.008) | (0.065) |
| University | | -0.0447*** | -0.0394*** | -0.0420*** | -0.0391*** | -0.0891 |
| | | (0.007) | (0.007) | (0.011) | (0.011) | (0.078) |
| Number of positions advertised | | 0.4817*** | 0.4492*** | 0.5448*** | 0.7139*** | -0.0909 |
| | | (0.087) | (0.086) | (0.136) | (0.153) | (0.583) |
| Occupation Fixed Effects | | | Y | Y | Y | Y |
| Job Title Fixed Effects | | | | Y | Y | |
| Firm Fixed Effects | | | | | Y | |
| Title*Firm Fixed Effects | | | | | | Y |
| "Effective" $N$ | 42,744 | 42,744 | 42,744 | 25,438 | 23,819 | 1,448 |
| $R^2$ | 0.010 | 0.434 | 0.437 | 0.531 | 0.738 | 0.846 |

Standard errors in parentheses, clustered by firm. *** p<0.01, ** p<0.05, * p<0.1

Notes:

1. Dependent Variable: the share of applicants that satisfy the job's age requirement, mean = .8532.

2. Regressions are weighted by the number of applications to the ad.

Table A.16: Effects of Gender Requests on the Share of Applicants satisfying the Job's Education, Experience and Age Requirements

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Ad requests men ($M$) | -0.0217*** | -0.0059 | -0.0056 | -0.0015 | -0.0003 | 0.0251 |
| | (0.006) | (0.004) | (0.004) | (0.006) | (0.006) | (0.048) |
| Ad requests women ($F$) | -0.0329*** | -0.0342*** | -0.0339*** | -0.0162*** | -0.0145** | -0.0256 |
| | (0.006) | (0.004) | (0.004) | (0.006) | (0.006) | (0.029) |
| Primary School | | 0.0426*** | 0.0438*** | 0.0413*** | 0.0481*** | 0.0440 |
| | | (0.006) | (0.006) | (0.008) | (0.009) | (0.046) |
| Middle School | | 0.0661*** | 0.0675*** | 0.0513*** | 0.0582*** | -0.0357 |
| | | (0.008) | (0.008) | (0.010) | (0.012) | (0.075) |
| Tech School | | -0.0242*** | -0.0238*** | -0.0156** | -0.0091 | -0.0109 |
| | | (0.006) | (0.006) | (0.008) | (0.008) | (0.041) |
| Post-secondary | | -0.1256*** | -0.1259*** | -0.1336*** | -0.1284*** | -0.1262** |
| | | (0.005) | (0.005) | (0.008) | (0.008) | (0.056) |
| University | | -0.3042*** | -0.3091*** | -0.3510*** | -0.3554*** | -0.3411*** |
| | | (0.006) | (0.007) | (0.010) | (0.011) | (0.073) |
| Number of positions advertised | | 0.2597*** | 0.2772*** | 0.2973** | 0.3770** | 0.1473 |
| | | (0.090) | (0.089) | (0.126) | (0.155) | (0.753) |
| Occupation Fixed Effects | | | Y | Y | Y | Y |
| Job Title Fixed Effects | | | | Y | Y | |
| Firm Fixed Effects | | | | | Y | |
| Title*Firm Fixed Effects | | | | | | Y |
| "Effective" $N$ | 42,744 | 42,744 | 42,744 | 25,438 | 23,819 | 1,448 |
| $R^2$ | 0.003 | 0.58 | 0.584 | 0.672 | 0.802 | 0.879 |

Standard errors in parentheses, clustered by firm. *** p<0.01, ** p<0.05, * p<0.1

Notes:

1. Dependent Variable: share of applicants that satisfy the job's education, experience and age requirements, mean = .7113.
2. Regressions are weighted by the number of applications to the ad.

## A.6 Modeling Implicit 'Maleness' and 'Femaleness' of Job Titles: A Naïve Bayes Approach

This note describes how we construct a measure of the perceived, or 'implicit' maleness of each job title using a Naïve Bayes approach based on the words in all the job titles. The same method can be used to construct job titles' 'implicit' femaleness. Our approach follows the algorithm described in Mitchell (1997). More specifically, such algorithm, which is commonly used in textual analysis, is referred to as the multivariate Bernoulli event model by McCallum et al. (1998).

### A.6.1 Description of the Problem

Let $J$ be the set of jobs, $K$ be the set of job titles that ever appear in the job set $J$, and $W$ be the set of words that ever appear in the job title set $K$. Define $|A|$ to be the number of elements in set A. Similarly, $|J|$ is the number of jobs, $|K|$ is the number of unique job titles and $|W|$ is the number of unique words in the job titles.

For any job $j \in J$, let $k(j) \in K$ be its title, and let $\omega(j) \in 0, 1$ indicate whether this job explicitly prefer men. In other words, $\omega(j) = 1$ if this job explicitly prefers men, and 0 otherwise. For any job title $k \in K$, let $W^k \subseteq W$ be the set of words that appear in job title $k$.

The implicit maleness of a job title $k$ with word set $W^k$ can then be expressed using Bayes rule as follows,

$$P(\omega = 1|W^k) = \frac{P(W^k|\omega = 1) \cdot P(\omega = 1)}{P(W^k)} \tag{A.1}$$

### A.6.2 Solving the Problem

Notice that $P(\omega = 1|W^k)$ can be rewritten as follows,

$$P(\omega = 1|W^k) = \frac{1}{1 + \frac{P(W^k|\omega=0)\cdot P(\omega=0)}{P(W^k|\omega=1)\cdot P(\omega=1)}} \tag{A.2}$$

#### The Prior Probabilities

One option for modelling the prior probabilities $P(\omega = 1)$ and $P(\omega = 0)$ is to use the overall share of jobs that explicitly prefer men and that of jobs that do not explicitly prefer men in the sample. This approach is indeed widely used in commonly text classification. While this information is available to us, it may not be available to individual jobseekers, whose perceptions we are attempting to model. Thus we adopt the naïve assumption that $P(\omega \hat{=} 1) = P(\omega \hat{=} 0) = 0.5$. Graham (2002) also argues for this assumption in the spam-filtering setting. Thus, equation A.2 simplifies to

$$P(\omega \hat{=} 1|W^k) = \frac{1}{1 + \frac{P(W^k|\omega=0)}{P(W^k|\omega=1)}} \tag{A.3}$$

#### The Conditional Probabilities: From Words to Job Titles

To simplify the challenging task of estimating $P(W^k|\omega)$, the Naïve Bayes approach assumes,

(1) the appearance of each word is independent, and

(2) the ordering of the words in a job title is irrelevant.

This implies that

$$P(W^k|\omega = 1) = \prod_{w \in W^K} P(w|\omega = 1) \tag{A.4}$$

and

$$P(W^k|\omega = 0) = \prod_{w \in W^K} P(w|\omega = 0) \tag{A.5}$$

**Estimation of Each Word's Conditional Probability**

For the estimation of $P(w|\omega)$, if we have a large enough sample we can use

$$a \cdot P(\hat{w}|\omega) = P(\omega|w) = \frac{|\{j : j \in J, w \in W^{k(j)}, \omega(j) = \omega\}|}{|\{j : j \in J, w \in W^{k(j)}\}|} \tag{A.6}$$

where $a \equiv \frac{P(w)}{P(\omega)}$ is assumed to be a constant and cancels out in the division of A.3.

In practice, however, even large samples frequently yield zeros in A6.4. Given equations A6.3a and A6.3b, we would then get zeros for the entire job title regardless of the other words in the title. To avoid this problem, we use a weighted average of $P(\hat{w}|\omega)$ and a constant number close to one as our estimate of $P(w|\omega)$. The formula is

$$P(\tilde{w}|\omega) = \frac{|\{j : j \in J, w \in W^{k(j)}\}|}{|\{j : j \in J, w \in W^{k(j)}\}| + C} = \frac{C}{|\{j : j \in J, w \in W^{k(j)}\}| + C} \cdot \frac{C-1}{C} \tag{A.7}$$

Furthermore, notice it is particularly important to adjust the $P(\hat{w}|\omega)$'s when the total number of $|\{j : j \in J, w \in W^{k(j)}\}|$ is small. That is, we do not want to have a linear adjustment. Instead, we want to pull $P(\hat{w}|\omega)$ towards $\frac{C-1}{C}$ more strongly the less frequently a word appears in job titles.

In the literature, the recommended value of $C$ is $|W|$. For maleness, $\frac{1}{|W|}\sum_{w \in W} P(\hat{w}|\omega) \approx 0.212$, $\frac{1}{|W|} = \frac{1}{5954} = 0.00017$. If we were to use $|W|$ as $C$, $P(\tilde{w}|\omega)$ would be substantially higher than $P(\tilde{w}|\omega)$ for most words. Therefore, to keep the distortion to a minimum, we choose $C$ to be the average number of $|\{j : j \in J, w \in W^{k(j)}\}| \approx 15.04$. Combining (A.4) and (A.5), we can get $P(\tilde{w}|\omega)$ as presented in (A.7).

To sum up, our estimator for the implicit maleness of a job title $k$ is

$$P(\omega \hat{=} 1|W^K) = \frac{1}{1 + e^{f(\omega=1|W^K)}} \tag{A.8}$$

where

$$f(\omega = 1|W^K) = \sum_{w \in W^k} \{ln[1 - P(w|\tilde{\omega} = 1)] - lnP(w|\tilde{\omega} = 1)\} \tag{A.9}$$

$$P(w|\tilde{\omega} = 1)\frac{|\{j : j \in J, w \in W^{k(j)}, \omega(j) = \omega\}| + C - 1}{|\{j : j \in J, w \in W^{k(j)}\}| + C} \tag{A.10}$$

where $C = \frac{1}{|W|}\sum_{w \in W} |\{j : j \in J, w \in W^{k(j)}\}|$.

### A.6.3 Final remarks

This note has described our machine-learning approach to estimating the likelihood that a job title will explicitly request men (or women) based on the words contained in the title. Notably, the purpose of our approach differs from the usual application of document classification algorithms, which in this case would be to produce the best possible forecast of the gender label an employer will attach to a job from all the data available to us. Instead we seek to model the perceptions of individual job-seekers who have less information than us, and who face time constraints and limited cognitive capacity. Thus we have adopted a relatively simple approach with a naïve prior, and abstained from elements that would be considered in an industrial textual analysis setting, such as a more detailed tokenization of words, dropping less frequent words, or using a term frequency-inverse document frequency (TF-IDF) approach to identify the more informative words in each job title.

## A.7 Gender Misclassification

Miscoding of the requested gender is not a concern for our application analysis, since our data are the exact record of requested gender that workers observe on the job board when deciding where to apply. Miscoding of the requested gender could account for the relatively high success rates of gender-mismatched applicants if employers sometimes specify a gender requirement without intending to. If so, advertised gender requirements would be de facto rather soft. We view this as a possible interpretation of the relatively weak mismatch penalty in callbacks in our data.

Another possibility is that workers miscode their own gender when using the drop-down menu in the application process. The very high compliance rates we observe suggest that this is not a major concern. Nevertheless, we checked to see if miscoded applicant gender could account for the relatively weak enforcement in our data by re-running the main analysis on a restricted subsample for whom we are confident we have the right gender.[9]

To construct this sample, we first use the universe of applications, with no restrictions, to calculate the share of applications each CV in the sample sends to jobs which request the opposite gender. We then drop all the CVs in our sample for whom this share is 0.5 or higher. We also drop all CVs who submit fewer than 5 applications in the unrestricted data, because we may not have enough observations on them to reliably assess their application behavior. These restrictions only drop approximately 15,000 applications, leaving a sample size of 213,719.

We then re-run the application-level regressions from Table 1.4, and the results are

[9] Note that miscoded applicant gender cannot explain weak enforcement if firms use resume-processing software to pre-screen resumes based on coded gender: such screens would eliminate both actual and false gender mismatches from consideration, generating a high level of measured enforcement. Miscoded applicant gender can only explain low compliance if employers can see that some apparently mismatched applicants are in fact of the requested gender (for example from the photo, name or other features of the resume.

very similar to those presented in the main analysis, which gives us confidence that the results are not being driven by misreported gender. They are reported in Table A.17. Results for other cutoffs are not materially different.

Table A.17: Effects of Job Labels (F, N and M) on Callback Rates for Gender Misclassification Robust Sub-Sample

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Female Worker * Female Job | -0.0140 | -0.0094*** | -0.0090*** | -0.0092*** | -0.0132*** | -0.0155*** |
| | (0.009) | (0.002) | (0.002) | (0.002) | (0.002) | (0.003) |
| Female Worker * Male Job | -0.0429*** | -0.0416*** | -0.0413*** | -0.0401*** | -0.0334*** | -0.0365*** |
| | (0.013) | (0.004) | (0.004) | (0.005) | (0.006) | (0.008) |
| Male Worker * Female Job | -0.0341*** | -0.0288*** | -0.0288*** | -0.0229*** | -0.0226*** | -0.0218*** |
| | (0.010) | (0.003) | (0.003) | (0.004) | (0.004) | (0.005) |
| Male Worker * Male Job | 0.0044 | 0.0012 | 0.0014 | 0.0031 | -0.0057 | -0.0157*** |
| | (0.009) | (0.002) | (0.002) | (0.002) | (0.004) | (0.005) |
| Male Worker | 0.0036 | 0.0007 | -0.0022 | -0.0064*** | -0.0166*** | |
| | (0.006) | (0.002) | (0.002) | (0.002) | (0.002) | |
| Education less than requested | | -0.0066*** | -0.0060** | -0.0080*** | -0.0087*** | -0.0093*** |
| | | (0.002) | (0.003) | (0.003) | (0.002) | (0.004) |
| Education more than requested | | -0.0041*** | -0.0075*** | -0.0063*** | -0.0017 | 0.0017 |
| | | (0.001) | (0.002) | (0.002) | (0.002) | (0.003) |
| Age less than requested | | -0.0008 | -0.0020 | -0.0023 | -0.0044** | -0.0022 |
| | | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| Age more than requested | | -0.0320*** | -0.0301*** | -0.0279*** | -0.0207*** | -0.0216*** |
| | | (0.003) | (0.003) | (0.003) | (0.003) | (0.004) |
| Experience less than requested | | -0.0059*** | -0.0062*** | -0.0076*** | -0.0093*** | -0.0073*** |
| | | (0.002) | (0.002) | (0.002) | (0.002) | (0.003) |
| Experience more than requested | | 0.0005 | 0.0017 | 0.0012 | -0.0012 | 0.0014 |
| | | (0.002) | (0.002) | (0.002) | (0.002) | (0.004) |
| Wage below advertised | | -0.0021 | -0.0020 | -0.0030 | -0.0003 | -0.0007 |
| | | (0.002) | (0.002) | (0.002) | (0.002) | (0.003) |
| Wage above advertised | | 0.0011 | 0.0008 | 0.0002 | -0.0057** | -0.0045 |
| | | (0.002) | (0.002) | (0.002) | (0.002) | (0.003) |
| Detailed CV controls | | | Y | Y | Y | |
| Occupation Fixed Effects | | | | Y | Y | Y |
| Competition Controls | | | | | Y | Y |
| Job Title Fixed Effects | | | | | Y | Y |
| Worker Fixed Effects | | | | | | Y |
| 'Effective' $N$ | 213,719 | 213,719 | 213,719 | 213,719 | 213,676 | 189,485 |
| $R^2$ | 0.001 | 0.004 | 0.005 | 0.015 | 0.194 | 0.383 |

# A.8 Implications for Gender Segregation

## A.8.1 Background

The goal of this Appendix is to illustrate the mechanisms via which a gendered-ad ban might affect labor market segregation, and to assess the implications of our findings for the size of those effects. Our approach is based on the idea that prohibiting explicit gender requests removes a piece of information that directs workers' applications away from jobs requesting the other gender; thus a gendered-ad ban will result in more gender-mismatched applications.[10] The effects of a gendered ad ban therefore depend on (a) the number of applications that are redirected, and (b) how those redirected applications are treated by employers. We calculate (a) using our regression estimates of female applicant shares ($\alpha$) from column 6 of Table 1.3, and – in our baseline calculations – we assume (b) is unchanged by a gendered ad ban. In other words, if (for example) female applicants to explicitly male jobs were 44.5 percent as likely to get a callback as men when gendered ads were allowed ($\theta = .445$ in Table 1.1), we assume those same jobs (which are no longer explicitly labeled as male) will continue to call back women and men in the same proportion after such ads are banned.

We proceed in four stages. First, we estimate the total amount of gender segregation among successful applications, (i.e. among called-back workers) at three different levels: the job (i.e. the ad), the firm, and the occupation. Next, we decompose these segregation measures into segregation within versus between the three job types defined by the explicit labels (F, N and M), and assume that within-label segregation is not affected by an ad ban: removing, say, the female label does not have any obvious ef-

[10] In all these exercises, we classify jobs according to their gender request before the ban. For simplicity, our approach holds constant the total number of applications and callbacks made at every job; only their gender composition is changed. Thus we abstract from any equilibrium changes in search and recruiting intensity that might be caused by a ban.

fects on how workers will choose among the jobs that were formerly labeled as female. Third, we simulate between-label post-ban segregation using the regression estimates from Table 1.3. Adding this counterfactual between-label segregation to within-label segregation gives us total segregation after a ban. Finally, we assess the robustness of our calculations to changes in assumptions.

### A.8.2 Measuring Segregation

To measure segregation, we use Duncan and Duncan (1955) segregation index, applied to the set of successful applicants (i.e. callbacks) in a unit (job ad, firm, or occupation). Duncan and Duncan's index, $S$, is calculated as:

$$S = \frac{\sum_i Y_i |\delta_i - \Delta|}{2\Delta(1-\Delta)} \tag{A.11}$$

where $\delta_i$ is the female share of callbacks in unit $i$, $\Delta$ is the female share in the population, and $Y_i$ is unit $i$'s share of the callback population. Thus, $S$ is the population-weighted mean absolute deviation of the female share from its global mean, divided by its maximum attainable value, $2\Delta(1-\Delta)$.[11] Like our gender matching index, Duncan and Duncan (1955) $S$ index varies between 0 and 1. It is widely used in studies of residential segregation (Cutler et al., 1999; Logan et al., 2004). In our context, $S$ gives the share of men (or women) who would have to be reassigned to a different occupation, firm or job in order for men and women to be distributed identically across those categories.[12]

[11] Equivalently, S can be calculated via the better known formula, $S = \frac{1}{2}\sum_j [\frac{\phi_i}{\Phi} - \frac{\mu_i}{M}]$, where $\phi_i$ is the share of callbacks in unit $i$ that go to women, $\mu_i = 1 - \phi_i$ is the share of callbacks in unit $i$ that go to men, and $\Phi$ and $= 1 - \Phi$ are their population equivalents.

[12] This property is independent of which group is being re-allocated and of the relative size of the two groups (Zoloth, 1976). Notably, however, the counterfactual reallocation of residents underlying this interpretation does not preserve the total populations of the units.

Because some of the units (especially jobs) used in our analysis are small, however, we need to adjust Duncan and Duncan's measure for the effect of purely random variation in where workers send their resumes and in which resumes are picked from the application pool.[13] To accomplish this, we extend the one-stage sample-shuffling approach developed by Carrington and Troske (1997) to reflect the fact that the allocation of workers to jobs is the outcome of two urn-ball processes: the allocation of applicants to jobs, and the selection of successful applicants from applicant pools.

In more detail, Carrington and Troske (1997) estimated the amount of racial segregation across Chicago workplaces we would expect if we took as given total employment at each workplace, and then imagined that the actual population at each workplace was a random draw from a binomial distribution whose mean black share was the population average. Simulating the Duncan-and-Duncan segregation index over multiple replications, then taking the mean of the resulting indices gave them an estimate of the amount of segregation we would see if workers were allocated to jobs in a race-blind way. In our context, we take as given the total number of applications and callbacks at every job ad. We then simulate the amount of segregation we would expect if the gender mix of applications to each ad, and of callbacks to each ad was the result of a random draw from binomial distributions with parameters derived from the population mean levels of $\alpha$ and $\theta$ in Table 1.1. The idea is to hold fixed the total number of applications men and women make, the number of applications arriving at each job, and the total number of 'interview slots' (callbacks) available for each job. With these 'structural' features of the labor market fixed, we then assume that workers direct their applications randomly and that firms select candidates randomly. How much gender segregation would we expect to see?

[13] This is especially important when measuring segregation across individual job ads, whose callback pools contain an average of 5.3 workers. To see the issue, note that if each ad calls back only one worker, segregation will always be complete: every job's callback pool will be entirely male or entirely female.

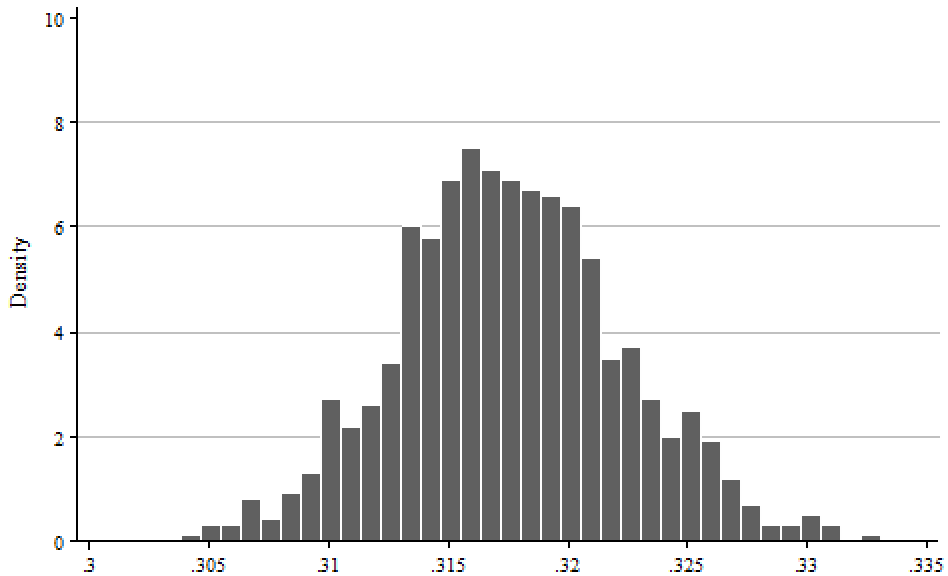
To illustrate, recall that the overall mean of $\alpha$, $\bar{\alpha} = .541$ and consider an ad that received 80 applications and issued 5 callbacks. We first simulate the number of female and male applications to that ad ($a^f$ and $a^m$) as a random draw of 80 applications from a pool with population parameter .541, i.e. $a^f \sim B(n,p) = B(80, .541)$, $a^m = 80 - a^f$, and $B$ indicates the binomial distribution. Next, taking this randomly-generated application pool as given (say, 51 women and 29 men), we simulate the number of male and female callbacks ($c^f$ and $c^m$) as a random draw of 5 callbacks from a pool with population parameter given by:

$$p^c = \frac{\bar{\theta} a^f}{\bar{\theta} a^f + a^m} \tag{A.12}$$

where $\bar{\theta} = .866$ is the overall mean of women's relative callback risk. Thus, $c^f \sim B(n,p) = B(5, p^c); c^m = 5 - c^f$. Doing this for every job, then calculating the realized segregation index, $S$, completes a single iteration.

Figure A.5 plots the distribution of realized $S$ values from 1000 iterations in this baseline scenario where there is no systematic variation across jobs in either application or callback behavior. It shows a surprisingly concentrated distribution with a mean of .317 and all values falling between .30 and .34. Thus, while random matching can generate a high level of measured segregation, the amount of segregation it generates is tightly constrained by the distribution of applicant pool sizes and callback pool sizes and the overall share of men and women in the population.

Figure A.5: Simulated Segregation Indices with Random Allocation of Applications to Jobs, and Random Selection of Callbacks from All Applicant Pools




Finally, to remove the effects of this randomness, we follow Carrington-Troske by defining a noise-adjusted segregation measure, $\tilde{S}$, as:

$$\tilde{S} = \frac{S - S_0}{1 - S_0} \tag{A.13}$$

where $S$ is the unadjusted segregation index from equation (A8.1) and $S_0 = .317$ is the mean level of segregation expected from noise in matching. Since $S = .732$, the noise-adjusted index of gender segregation across jobs in our data is given by $\tilde{S} = \frac{.732 - .31}{1 - .317} = .607$.

Unadjusted $(S)$ and noise-adjusted $(\tilde{S})$ Duncan segregation indices across jobs, firms and occupations are shown in columns 1 and 2 of Table A.18. Interestingly, noise-adjusted segregation across jobs equals .607, which essentially coincides with Cutler et al. (1999) threshold of 0.6 for defining a U.S city as having a residential ghetto. Adjusted segregation across other units is lower (at .394 and .385 for firms and occupations respectively), and – as expected – adjusting for random matching has the greatest impact in the smallest units (jobs).

### A.8.3 Decomposing Segregation

We begin by noting that a substantial amount of the gender segregation among successful applicants occurs within groups of jobs that have a specific gender label (F, N or M) attached to them; this component of gender segregation is unlikely to be impacted by banning the labels. To calculate it, we first calculate total (unadjusted) between-label segregation $S^B$ by simulating the amount of segregation that would exist if each of the three explicit job types had its own $\alpha$ and its own $\theta$ (given by the raw means in the data), but all remaining allocation of workers to jobs and callbacks to workers was random. Adjusting this for noise yields the noise-unadjusted amount of between-label segregation:

$$\tilde{S}^B = \frac{S^B - S_0}{1 - S_0} \tag{A.14}$$

The amount of within-label segregation is then given by subtraction:

$$\tilde{S}^W = \tilde{S} - \tilde{S}^B \tag{A.15}$$

As reported in columns 3 and 4 of Table A.18, noise-adjusted between- and within-label gender segregation across jobs equal .360 and .247 respectively; thus almost 60 percent of overall segregation (.607) is between groups of jobs defined by their explicit gender requests. In the remainder of this Appendix we compute the predicted effects of a gendered ad ban on between-label segregation using our regression estimates. Assuming that within-label segregation is not affected by an ad ban, we then compute the percentage decline in overall segregation that would be caused by a ban under a variety of assumptions.

Table A.18: Actual and Segregation across Job Titles, Occupations and Firms

| Gender Segregation across | Unadjusted Segregation | Noise-adjusted Segregation | | | | | |
|---|---|---|---|---|---|---|---|
| | Total | Total | Between-label | Within-label | Between label (non-causal) | After a gendered-ad ban | Reduction from a gendered ad ban (%) |
| | $S$ | $\tilde{S}$ | $\tilde{S}^B$ | $\tilde{S}^W = \tilde{S} - \tilde{S}^B$ | $\tilde{S}^{BNC}$ | $\tilde{S}^A = \tilde{S}^W + \tilde{S}^{BNC}$ | $\frac{\tilde{S} - \tilde{S}^A}{\tilde{S}}$ |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Jobs | 0.732 | 0.607 | 0.360 | 0.247 | 0.190 | 0.438 | 0.279 |
| Firms | 0.505 | 0.394 | 0.234 | 0.160 | 0.126 | 0.287 | 0.273 |
| Occupations | 0.405 | 0.385 | 0.204 | 0.181 | 0.131 | 0.312 | 0.189 |

### A.8.4 Baseline Effects of a Gendered-Ad Ban

To estimate between-label segregation after a ban, we proceed in two stages. First, in Table A.19 we use the regression coefficients from column 6 of Table 1.3 to estimate the female applicant shares to F, N and M jobs that we would expect to see after a ban: these reflect only the differences in $\alpha$ across job types that are not caused by the ban. (Men and women will continue to apply to typically-male and -female jobs even after a ban, just not to the same extent as before.) Second, we use these applicant shares to simulate between-label segregation after a ban, assuming – in our baseline scenario – that the additional applications to gender-mismatched jobs continue to be treated the same by employers (i.e. encounter the same $\theta s$) as they were before the ban. We then add this counterfactual between-label segregation to within-label segregation to get our baseline estimate of segregation after a ban.

#### Female applicant shares after a ban

When gendered ads are permitted (the situation for which we have data), the female shares of applications to F, N and M jobs are .926, .447 and .079 respectively. These shares, from row 2 in Table 1.1, are reproduced in row 1 of Table A.19. According to our estimates, how are these shares likely to change when gendered ads are banned? To estimate this, we partition the raw compliance effects (.479 = .926 - .447 and -.368 = .079 - .447) into two components: a causal effect of the explicit gender requests (given by column 6 of Table 1.3: .246 and -.146 respectively) and their complement (the non-causal components: .233 and -.222 respectively). Since banning gendered ads removes only the causal component of gendered ads – the part that directs workers' applications – we then calculate a counterfactual set of $\alpha$'s that reflects only the non-causal component, reported in row 5 of Table A.19. Notice that even when gendered ads are prohibited, our estimates imply that women will still disproportionately apply to the (often stereotypically female) jobs that formerly requested women, just less disproportionately than before the ban.

**Between-label and overall segregation after a ban**

Given that a gendered ad ban will increase the number of applications to (formerly) gender-mismatched jobs (Row 5 of Table A.19 is less differentiated than row 1), the effects of a gendered ad ban on gender segregation depends on how those new, mismatched applications are treated by employers. While it is conceivable that these applications could be treated either more or less harshly than mismatched applications before the ban, in our baseline calculations we suppose they are treated in the same way. In other words, we shall assume that the relative callback rates of men and women in F, N and M jobs (the $\theta s$) are just given by the sample means in our data, shown in row 5 of Table 1.1.

Using this assumption, we now simulate between-label segregation after a ban using the applicant shares in row 5 of Table A.19 to. For job segregation this equals .190, as reported in column 5 of Table A.18. We then add this post-ban between-label segregation to within-label segregation to get our baseline estimate of segregation after a ban, reported in column 6 of Table A.18. According to those estimates, a gendered ad ban is predicted to reduce gender segregation across jobs, firms and occupations by 27.9, 27.3 and 18.9 percent respectively in this baseline case.

Table A.19: Actual and Counterfactual Female Share of Applications ($\alpha$)

| | Ad Requests Women *F* jobs | Gender not specified *N* jobs | Ad Requests Men *M* jobs | All Ads |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| 1. Actual share of applications that are female ($\alpha$) (Baseline) | 0.926 | 0.447 | 0.079 | 0.541 |
| 2. Raw compliance effects | 0.479 | | -0.368 | |
| 3. Causal effect of gender requests | 0.246 | | -0.146 | |
| 4. Non-causal component (difference between row 2 and 3) | 0.233 | | -0.222 | |
| 5. Estimated female share of applications after an ad ban ($\alpha^A$) (reflects the non-causal component of gender labels only). | 0.734 | 0.501 | 0.279 | 0.541 |

Notes:

1. Row 1 is from row 2 of Table 1.1; row 3 is from column 6 of Table 1.3.
2. The raw compliance effect of row 2 is calculated as the difference between column 1 and 2 of row 1 for female jobs and column 3 and 2 of row 1 for male jobs.
3. Row 4 is calculated as the difference between row 2 and 3.
4. The female applicant shares in row 5 are calculated to reflect the non-causal female-share differences between job types in row 4 (for example, .734 - .501 = .233), while preserving the grand mean of $\alpha$ across all ads (.541).

While these declines are substantial, we note that the magnitude of these declines is constrained by two key features of our simulations: (a) we do not expect an ad ban to change the amount of within-label segregation; and (b) our Table 1.3 regressions indicate that many workers will continue to disproportionately apply to jobs that formerly requested their gender even after a ban, most likely because they have tastes and training that attach them to those types of work. Indeed, gender-specific training may help to explain why ad bans have a larger predicted effect at the job title and firm level than the occupation level: It may be easier to re-allocate one's application to a different firm or different detailed job title in response to the removal of explicit gender requests than to change one's occupation.[14] This finding is also consistent with our result (in Section 3.2) that gender labels affect application behavior most strongly in jobs which are not clearly gender-stereotyped – it is in these low-$F_p$ and low-$M_p$ jobs where we expect the desegregating effects of an ad ban to be the greatest.

### A.8.5 Sensitivity and Caveats

Is it realistic to assume that the callback penalties faced by gender-mismatched applicants will be unaffected by a gendered ad ban? To explore this issue, we now comment on why a ban might cause mismatched applicants to be treated either more or less harshly, and estimate the effects of a gendered ad ban under some alternative assumptions. The specific cases we examine are summarized in Table A.20: a 50 percent increase in both men's and women's mismatch penalties after the ban, and an elimination of callback penalties after the ban.

[14] Mechanically, gender differences in occupation-specific training affect our simulation results via two channels: First, compared to job segregation, a larger share of occupational segregation is within the F, N, and M labels, and therefore not affected by a ban. In particular, the majority of jobs that are not explicitly gendered (N jobs) are highly gender segregated by occupation, and this segregation is unlikely to be changed by removing gender requests. Second, again compared to job segregation, a smaller share of between-label occupational segregation is caused by the labels.

Table A.20: Alternative Assumptions about relative callback rates ($\theta$) after a gendered-ad ban

| | **Ad Requests Women** | **Gender not specified** | **Ad Requests Men** |
|---|---|---|---|
| | ***F* jobs** | ***N* jobs** | ***M* jobs** |
| 1. Baseline (row 5, Table 1) | 1.246 | 0.958 | 0.445 |
| 2. Mismatch penalties increase by 50% | 1.741 | 0.958 | 0.216 |
| 3. Mismatch penalties eliminated | 0.958 | 0.958 | 0.958 |

Notes:

Row 2 of the Table multiplies the mismatch penalties implied by rows 3 and 4 of Table 1.1 by 1.5, then calculates the resulting $\theta s$ as the ratios of the new callback rates.

**Increased differentials in $\theta$ across job types.**

Suppose that in the presence of explicitly gendered job ads, only very highly qualified workers applied to gender-mismatched jobs. If this was the case, banning gendered ads could create a new batch of gender-mismatched applications that are less qualified than before. In this case, we would expect mismatched applications to be treated, on average, more harshly after a ban than before. To explore the implications of this effect, Panel B of Table A.21 replicates Table A.18 for the case where both men's and women's mismatch penalties are 50% larger in magnitude after the ban (row 2 of Table A.20). We find that a gendered ad ban still reduces gender segregation (because it redirects applicants to gender-atypical jobs), but the predicted decline (ranging from 13 to 22 percent) is considerably more modest because those applications are now treated more harshly than before. We do not think this scenario is likely, however, because it is inconsistent with our regression results for callbacks, which find negative self-selection into gender-mismatched jobs when the labels are visible to applicants.

Table A.21: Simulated Effects of a Gendered Ad Ban under alternative assumptions about the effects of a ban on $\theta$

| **Segregation across** | **Noise-adjusted segregation ($\tilde{S}$)** | **Estimated noise-adjusted segregation after a gendered-ad ban** | **Percentage reduction in noise-adjusted segregation from a gendered ad ban** |
|---|---|---|---|
| | $\tilde{S}$ | $\tilde{S}^A$ | $\frac{\tilde{S}-\tilde{S}^A}{\tilde{S}}$ |
| | (1) | (2) | (3) |
| **A. Baseline: θ is unaffected by a gendered ad ban** | | | |
| Jobs | 0.607 | 0.438 | 0.279 |
| Firms | 0.395 | 0.287 | 0.273 |
| Occupations | 0.385 | 0.312 | 0.189 |
| **B. Gender-Mismatch Penalties Increase by 50%** | | | |
| Jobs | 0.607 | 0.476 | 0.216 |
| Firms | 0.395 | 0.313 | 0.208 |
| Occupations | 0.385 | 0.335 | 0.131 |
| **C. Gender-Mismatch Penalties fall to zero** | | | |
| Jobs | 0.607 | 0.395 | 0.349 |
| Firms | 0.395 | 0.260 | 0.341 |
| Occupations | 0.385 | 0.285 | 0.259 |

**Diminished differentials in $\theta$ across job types**

There are at least two reasons why a gendered-ad ban might cause gender-mismatched applicants to jobs that were formerly explicitly gendered to be treated less harshly than before. One is the possibility that these new applicants are more positively selected than before (because before the ban, mismatched applicants were negatively selected). Second, a gendered-ad ban could signal to employers that public policy has become less tolerant of gender discrimination in the applicant selection process as well as the advertising process. To explore this scenario, panel C of Table A.21 explores the extreme case where a gendered ad ban eliminates gender mismatch penalties for both men and women (row 3 of Table A.20). Here the predicted declines in segregation are larger than in the baseline case, but not dramatically so, ranging from 26 to 35 percent. This modest effect of completely eliminating gender discrimination in employers' callback decisions process underscores the dominant role of workers' self-selection decisions in accounting for gender-matching in labor markets, already noted in our discussion of the aggregate statistics.

In sum, even large changes to our assumptions about post-ban call-back rates have only modest effects on the ban's predicted effects on segregation. This is because (a) within-label segregation is not directly affected by a ban, (b) only about 60 percent of between-label segregation is caused by explicit gender requests, and (c)gender differentials in callback rates have small effects on outcomes when applicant pools are highly segregated (as they are in our simulations, even after a ban).

Some important caveats regarding the above calculations is that they do not incorporate steps employers might take to circumvent a gendered ad ban, nor do they incorporate longer-term changes in workers' human capital investment decisions that might be caused by a ban. For example, employers might respond to a ban communicating the information formerly conveyed in explicit gender requests via other signals, including code words.[15] If such responses are common and effective, a gendered ad

[15]Some job boards have also responded to ad bans by making it easier for recruiters filter resumes by gender, both within the applicant pool and when a recruiter is searching through resumes posted on the site. Note that, because workers will no longer know whether firms are engaged in this filtering, these tools will raise search frictions for workers compared to a labor market that allows gendered ads.

ban might reduce job segregation by much less than our baseline estimate of 28 percent. On the other hand, if a ban signals to workers that investments in gender-atypical skills are more likely to be rewarded, its longer-run effects might be considerably larger than our estimates. For example, a ban might induce more men to train as nurses and more women to train as electricians.

Finally, we remind the reader that these predicted effects on gender segregation come at the cost of increased labor market frictions: because a ban directs a substantial number of workers' applications into jobs that formerly requested the other gender –where by assumption they have a smaller chance of getting a callback– fewer total callbacks will result from the same total number of applications. Stepping outside our simulations, after a ban, both men and women might need to submit more applications, search longer, or reduce their reservation wages to find an acceptable job.

# Appendix B

# Appendix for "Should I Show or Should I Hide – When Do Jobseekers Reveal Their Wages?"

## B.1 Model

We present the model in four stages below. First, we describe the three-stage structure of the model, then we solve Stage 3 of the model (employers' offer decisions) for the case where all workers must disclose their wages. Third, we solve Stage 3 when some workers choose not to disclose their wages. Finally, we characterize workers' wage disclosure decisions in Stage 1.

### B.1.1 Model Structure

In the model, the worker's current wage is the only indicator of a worker's quality, with a higher wage representing a higher level of productivity or ability. Worker $i$

$(i = 1, 2, \ldots, N)$ knows his current wage $w_i$, which remains private unless disclosed. If the worker discloses, he has no choice but to be truthful. Each worker applies to exactly $K$ jobs, which are randomly drawn from jobs posting a higher wage than his current wage. This assumption has two noteworthy consequences. First, the worker's search is *directed* in the sense that no workers apply to jobs paying less than the worker's current wage. This means that workers' application decisions are informative to firms. Second, this is the *only* sense in which workers' search is directed: workers are randomly matched to jobs paying more than their reservation wage. This allows us to think of jobs within that set as being randomly assigned to workers: Workers will learn the set of jobs they've been assigned to and make their disclosure decisions based on that set. Thus, whether a worker applies 'conservatively' or 'aggressively' is randomly assigned to the worker in the model.

There are $M$ employers (firms), and each of them posts one job position (so there are $M$ jobs in total). Labor is the only resource used in production, and higher-wage workers are more productive, implying that the production function $S(w)$ increases with $w$. Firms are differentiated in terms of production functions, $S^1, S^2, \ldots, S^M$, and higher indexed firms are more efficient in the sense that they can produce more output from workers with the same wage ($S^1(w) < \ldots < S^M(w)$, for any $w$). Employers receive job applications from workers, and make take-it-or-leave-it wage offers to some of their applicants. More specifically, employer $j$ commits to a wage offer function, $O^j(w)$, which assigns a wage offer to every potential applicant according to their current wage. Higher-wage workers achieve wage offers $(O'(w) > 0)$, and more efficient firms make higher wage offers to workers with the same current wage $(O^1(w) < \ldots < O^M(w)$, for all $w)$. Firm $j$'s profits from worker with wage $w$ $(R^j(w))$ are given by the difference between the worker's output and his wage: $R^j(w) = S^j(w) - O^j(w)$. Importantly, we assume that the offer function $O^j(w)$ *shares* the surplus from the employment relationship between workers and firms: Both profits and wage offers rise with the worker's quality, so firms are better off hiring abler workers and workers prefer to be hired by more productive firms.[1]

Knowledge about the labor market is distributed as follows. First, the workers' current wage distribution is public information, and (because of random matching) employers know the distribution from which the current wages of their applicants are drawn. Second, the distribution of firms' productivities is known to both employers and job seekers, and workers can recognize firms that are better (more efficient and can pay more) than their current employers. Third, employers know that each worker applies to $K$ jobs, but they do not know the wages of the other jobs their applicants have applied to. Finally, every employer computes its own expectation concerning the non-disclosing applicants it has received, and each worker knows his expected wages from all $K$ employers he has been assigned to.

Workers, firms and nature move in the following order:

In **Stage 1**, a worker decides either to disclose or withhold his wage ($D_i = 1$ if worker $i$ discloses). If the worker voluntarily discloses his current wage, his wage information is available to all the $K$ employers in his application set (i.e. who have received his resume).

In **Stage 2**, nature randomly assigns each worker to $K$ jobs, all of which offer a wage that (weakly) exceeds the worker's current wage.

In **Stage 3**, after receiving the above applications, employers make job offers to

[1]Empirically, of course, we only see whether a worker is marked as a recruiting target, not whether she receives a job offer, but since only 4.8 percent of applicants are marked as targets this seems like a reasonable approximation. A potentially less attractive (but not unusual) feature of this modeling strategy is that we assume all workers –including concealers– are offered the wage the firm has assigned to their type. Thus, firms who discover that a concealer had a lower wage than they expected may not revoke their wage offer.

workers by maximizing profits, given by:

$$\Pi(\mathbf{P}|\mathbf{W}, \mathbf{D}) = \theta_1 R(w_1)P_1 + \theta_2 R(w_2)P_2 + ... + \theta_n R(w_n)P_n - c\sum_{i=1}^{n} P_i \tag{B.1}$$

where the subscript $i$ indexes applicants in order of wages from high to low ($w_1$ is the highest wage). The wage of non-disclosers is equal to $W_E$. The offer decision is denoted by $\mathbf{P}(P_1, P_2, \ldots, P_n)$ where each element is binary, and $P_i = 1$ if the employer makes an offer to applicant $i$. Since employers need to pay a contact cost $c$ for each offer that is issued, the expected return to making an offer to a worker depends on the profit $R$ that the worker can generate, as well as the chance that the worker accepts the offer $\theta$. For instance, if employer $j$ has received $m_j$ job applications, the employer's expected return conditional on applicants' wages $\mathbf{W}(w_1, w_2, \ldots, w_{m_j})$ and their wage revealing decisions $\mathbf{D}(D_1, D_2, \ldots, D_{m_j})$ can be written as:

$$\mathbf{E}\Pi^j(\mathbf{P}|\mathbf{W}, \mathbf{D}) = \sum_{i=1}^{m_j} \mathbf{E}\theta_i^j(S^j(w_i) - O^j(w_i))P_i^j - P_i^j c_j. \tag{B.2}$$

Employer $j$ can make multiple callbacks, and makes an offer to the $i$th applicant ($P_i^j = 1$) as long as the expected return from applicant $i$, $R^j(w_i)\theta_i^j$ outweighs the contact cost, $c_j$.

In **Stage 4**, workers decide to accept or reject the offers they have received. Here we simply assume a worker can take only one job, so the worker's utility is given by the highest wage offer he has received.

Worker $i$'s utility, as a function of his wage and disclosure decision, and of the outcomes of his $K$ job applications can be written as follows:

$$U_i(D_i, \theta_i|w_i, \mathbf{P}) = max\{\theta_i^1 P_i^1 O_i^1, \theta_i^2 P_i^2 O_i^2, \ldots, \theta_i^K P_i^K O_i^K\} \tag{B.3}$$

where $\theta_i^j = 1$ if worker $i$ decides to accept the offer from job $j$, and all the other $\theta s$ are 0 ($\theta_i^k = 0, k = 1, \ldots, K, k \neq j$).

### B.1.2 Solving Stage 3 When All Wages are Revealed

To compute the chances that worker $i$ accepts an offer from firm $j$, we begin with the observation that worker $i$ accepts $j$'s offer only if employer $j$ proposes the highest wage among all the employers who make offers to $i$. Thus, by definition, firms that are better than firm $j$ do not extend offers to worker $i$, and the expected probability that a worker with wage $w$ will accept the offer from job $j$ is:

$$\begin{aligned}\mathbf{E}\theta^j|w &= Prob(\theta^j = 1) = Prob(max\{P^1O^1(w), \ldots, P^MO^M(w)\} \leq O^j|w) \\ &= Prob(P^{j+1} = 0, P^{j+2} = 0, \ldots, P^M = 0|w)\end{aligned} \tag{B.4}$$

using Bayes' Theorem,

$$
\begin{aligned}
\mathbf{E}\theta^j|w &= Prob(P^{j+1}=0|P^{j+2}=0,\dots,P^M=0|w)\cdot Prob(P^{j+2}=0,P^{j+3}=0,\dots,\\
&\quad P^M=0|w)\\
&= Prob(P^{j+1}=0|P^{j+2}=0,\dots,P^M=0|w)\cdot \mathbf{E}\theta^{j+1}|w\\
&= Prob(P^{j+1}=0|P^{j+2}=0,\dots|w)\cdot Prob(P^{j+2}=0|P^{j+3}=0,\dots|w)\\
&\quad \cdot Prob(P^{j+3}=0,\dots|w)\\
&= \dots\\
&= Prob(P^{j+1}=0|P^{j+2}=0,\dots|w)\cdot Prob(P^{j+2}=0|P^{j+3}=0,\dots|w)\cdot\\
&\quad Prob(P^{j+3}=0|P^{j+4}=0,\dots|w)\cdot\dots Prob(P^{M-1}=0|P^M=0|w)\cdot Prob(P^M=0)\\
&= \prod_{d=1}^{m-j} Prob(P^{j+d}=0|P^{j+d+1}=0,P^{j+d+2}=0,\dots,P^M=0|w)\\
&= q^{j+1}\cdot q^{j+2}\cdot\dots\cdot q^M
\end{aligned}
\tag{B.5}
$$

For an arbitrary firm $j+d$, if better firms $(j+d+1,\dots,M)$ do not make offers to a worker, the probability that the firm $j+d$ also rejects the worker monotonically decreases with the worker's wage $w$ (each term in the product, $q^{j+d}$, is a monotonically decreasing function of $w$). Then we can easily see that

$$
\begin{aligned}
\frac{d\mathbf{E}\theta^j|w}{dw} &= \frac{dq^{j+1}\cdot q^{j+2}\cdot\dots\cdot q^M}{dw}\\
&= \sum_{d=1}^{m-j}\frac{dq^{j+d}}{dw}q^{j+1}\dots q^{j+d-1}\dots q^{j+d+1}\dots q^M < 0
\end{aligned}
\tag{B.6}
$$

$\mathbf{E}\theta$ falls with the applicant's current wage $w$ (except for the best firm, $P^M$). The intuition is that a high applicant wage intensifies the competition among employers. This leads to a lower probability that the applicant will accept an offer from each employer, since there is a large chance that the high-wage worker has received better offers.[2] Moreover, for any given wage level, $\mathbf{E}\theta^j < \mathbf{E}\theta^{j+1}$ as workers always want to accept higher offers (offers from better firms). The expected return to employer $j$ from offering a job to applicant with wage $w$ can then be written as

$$\mathbf{E}\theta^j R^j|w = Prob(P^{j+1} = 0, P^{j+2} = 0, \ldots, P^M = 0|w)R^j(w). \tag{B.7}$$

The second component of equation (B.7) is the net revenues produced by an applicant with current wage $w$, which are increasing with the wage. The first term is the probability that applicant with wage $w$ accepts firm $j$'s the offer, which reflects the competition among employers and decreases with the wage. Taking these two points together, the probability a worker accepts an offer reflects the degree of match between the worker's productivity and the firm's productivity. This property leads to a form of positive assortative matching between workers and firms.

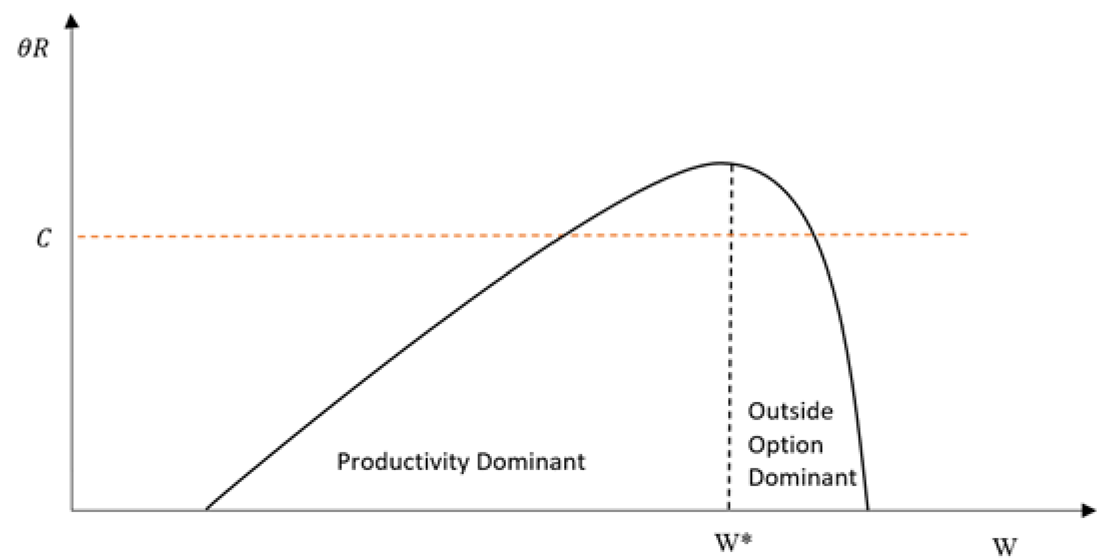
Figure B.1 illustrates how the expected return to making an offer changes with the worker's current wage for all employers except the most productive one.[3] When $w \leq w^*$, under regular conditions, $\mathbf{E}\theta R$ increases with the worker's wage $w$ because higher-wage workers produce more $(R(w))$. We call this the *productivity-dominant* case. However, when the applicant's current wage is too high (i.e. $w > w^*$), the applicant likely has offers from better firms that can offer higher wages. In this *outside option-dominant* case, the decrease in the probability the applicant accepts outweighs the increase in worker productivity and pulls down the firm's expected return from contact-

[2] Cahuc et al. (2006) find that between-firm competition matters a lot in the determination of wages for the on-the-job searchers.

[3] For the best employer $M$, the expected return is monotonically increasing with the worker's wage $w$.

ing this worker.[4]

Figure B.1: Expected Return of Making an Offer to Workers with Different Wages




With the expected profit function in hand, we can now describe the profit-maximizing offer policy *when the employer knows the current wages of all its applicants*: Employer $j$ will make an offer to a worker when the expected return exceeds the contact cost, $c_j$. According to Figure B.1, low-wage workers do not get offers because of their low productivity, whereas high-wage workers may be rejected due to their low chance of accepting an offer. Workers in a middle range of productivities will receive

[4]The expected return to employer $j$ from offering a job to applicant with wage $w$ is $\mathbf{E}\theta^j R^j|w$, and

$$\frac{d\mathbf{E}\theta^j R^j|w}{dw} = \frac{d\mathbf{E}\theta^j}{dw}R^j + \frac{dR^j}{dw}\mathbf{E}\theta^j \tag{B.8}$$

, which can be zero given $\frac{d\mathbf{E}\theta^j}{dw} < 0$ and $\frac{dR^j}{dw} > 0$ (to define a unique $w^*$). The second order condition

$$\frac{d^2\mathbf{E}\theta^j R^j|w}{d^2w} = \frac{d\mathbf{E}\theta^j}{dw}\frac{dR^j}{dw} + \frac{d^2\mathbf{E}\theta^j}{d^2w}R^j + \frac{d^2R^j}{d^2w}\mathbf{E}\theta^j + \frac{dR^j}{dw}\frac{d\mathbf{E}\theta^j}{dw} \tag{B.9}$$

can be negative under the sufficient conditions such that the marginal effect of wage on worker's offer acceptance probability is constant or diminishing (i.e., $\frac{d^2\mathbf{E}\theta^j}{d^2w} \leq 0$) and the marginal return on hiring abler worker is constant or decreasing (i.e.,$\frac{d^2R^j}{d^2w} \leq 0$). In this case, $\mathbf{E}\theta^j R^j|w$ is a concave function of $w$, as shown in Figure B.1.

offers.

### B.1.3 Solving Stage 3 When Some Wages are Concealed

To complete our description of an employer's optimal offer policy, we now need to specify how employers treat workers who have not disclosed their current wages. For simplicity, we assume that firm $j$ assigns an *expected* wage, $W_E^j$ to all its non-disclosing workers, and that these expected wages satisfy the following properties:

***Rule 1:*** Ordering of $W_E, W_E^1 < W_E^2 < \ldots < W_E^M$. Looking across firms, high-productivity firms are more optimistic about the current wages of their applicants. More specifically, the expected wages of non-disclosing applicants have the same ordering across workers as the corresponding employers' quality.[5] This ordering seems plausible given the assortative matching result described above.

***Rule 2:*** The expected wage of non-disclosing applicants is not higher than the expected-return-maximizing wage, i.e. $W_E \leq w^*$. Put another way, $W_E$ always falls in the productivity-dominant area. Since non-disclosing workers are choosing to hide their wages, it seems reasonable to assume that their wages are lower than the wage of the applicant that maximizes the firm's profits.[6]

### B.1.4 Workers' Optimal Disclosure Decisions in Stage 1

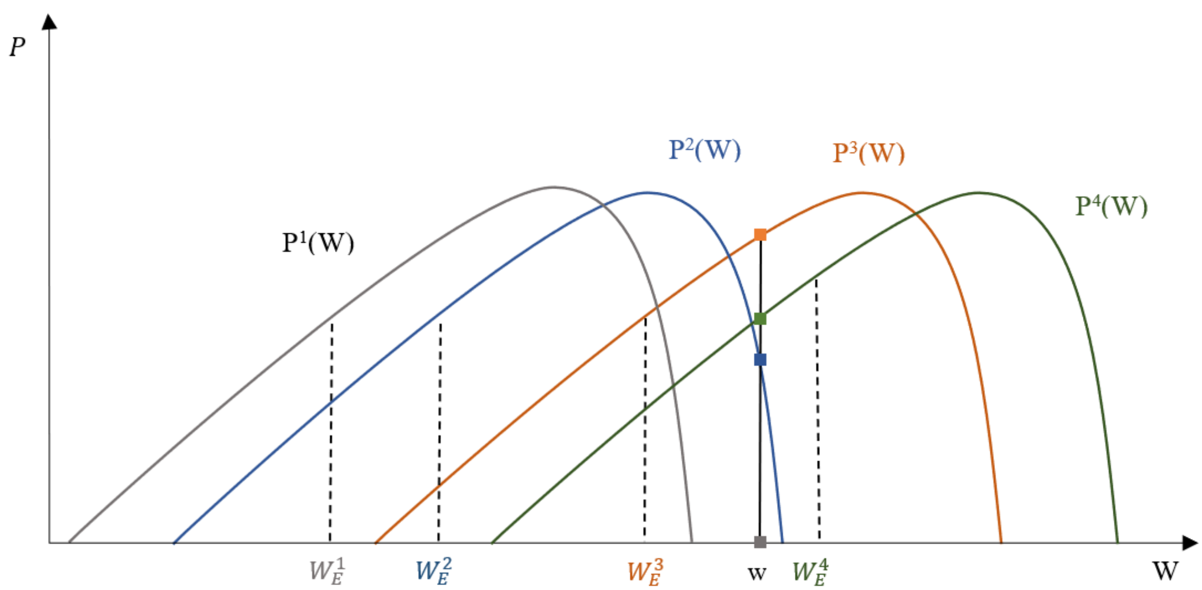
If the worker applies for multiple jobs ($K > 1$), the computation of worker's expected utility becomes more complicated because every employer will organize the

[5]For example, in one ordering that satisfies this property, all the workers apply to the best job, and the very top workers who only target the best firm will not apply to the second best firm, so the quality of the applicants for the second best firm is lower than the best firm, and this repeats to the lowest employer.

[6]For the best firm, $w_M^*$ is the highest wage of the applicants as $\theta$ is 1 and $R^M(w)$ increases with $w$, and the non-disclosed wages have to be lower than $w_M^*$. The second best firm has the same situation in which $w_{M-1}^*$ is the second highest wage of applicant and above the average wage of the applicants applying for the second best job. It continues until to the last employer.

expected wage of its non-disclosing applicants, leading to the other driver of the incomplete unraveling: When a non-disclosing worker is thought of having different wages in the perspective of different employers, he may face with the trade-off between disclosing wage to win a moderate callback and concealing wage to gamble for a callback from a superior firm.[7] The heterogeneity of the expected wages across employers motivates workers to manipulate their wage disclosure options when they apply for multiple jobs.

Figure B.2: Expected Offer Probability in Application Set




Next we turn to the wage disclosure decision of an individual worker. Suppose a worker with wage $w$ has applied $K$ jobs, and Figure B.2 demonstrates the expected callback probabilities of jobs in his application set ($K = 4$). Curves posit at higher wage levels represent more efficient employers, which attract better applicants with higher

[7]In the case of $K = 1$ where workers apply for only one job, the competition among employers would disappear given employers know workers' entire application set, $E\theta R$ therefore monotonically increases with wage (no outside-option dominated area), leading to the unraveling result (and a sorting model such that the $ith$ best worker matches the $ith$ best firm).

proposed wages and infer higher expected wages for the non-disclosing applicants. A worker's wage disclosure strategy depends on a comparison of the expected proposed wage from being evaluated by his actual wage, $U^D$, and that from being thought of holding the expected wage of each job, $U^C$. If $U^D \geq U^C$, the worker will choose to disclose his wage, otherwise he withholds.

$$U^D = max\{P^1(w)O^1(w), \ldots, P^V(w)O^V(w), \ldots, P^T(w)O^T(w), \ldots, P^K(w)O^K(w)\} \tag{B.10}$$

$$\begin{aligned} U^C = max\{&P^1(W_E^1)O^1(W_E^1), \ldots, P^V(W_E^V)O^V(W_E^V), \ldots, P^T(W_E^T)O^T(W_E^T), \\ &\ldots, P^K(W_E^K)O^K(W_E^K)\} \end{aligned} \tag{B.11}$$

Because employers infer the concealing applicants having the same wage $W_E$, $U^C$ is constant for workers who have applied to the same jobs. However, the value of $U^D$ depends on the worker's actual wage level, where each element can be decomposed into callback probability $P$ and proposed wage $O$. As noted above, proposed wage $O$ is increasing with respect to wage $w$, and for any wage level, we have $O^1(W) < O^2(W) < \ldots < O^K(w)$. A high wage increases the callback chance when the wage is in the productivity-dominant area, but has the opposite effect in the outside option-dominant area. To capture different trends of callback probabilities, we denote the best job in which the worker's wage $w$ is in the outside option-dominant area as job $V$ (i.e. $V = 2$ in Figure B.2), and $w$ is in the productivity-dominant areas for jobs $j > V$. Moreover, we define job $T$ as the best job having the expected wage that is lower than the worker's wage $w$, ($W_E^1 < W_E^2 < \ldots < W_E^T < W < W_E^{T+1} < \ldots W_E^K$) (i.e. $T = 3$ in Figure B.2). Since the expected wages are always in the productivity-dominant areas, $w$ must be higher than the expected wages of the first $V$ jobs, thus we have $V \leq T$.

For job $1$ to job $V$, $O^1(w) < O^2(w) < \ldots < O^V(w)$, and $P^1(w) < P^2(w) < \ldots < P^V(w)$ as $w$ falls in the outside option-dominant areas for all of the $V$ jobs. Thus $U^D_{(1,V)} \equiv max\{P^1(w)O^1(w), \ldots, P^V(w)O^V(w)\} = P^V(w)O^V(w)$. While the actual wage $w$ is higher than all the expected wages in this range, we cannot tell whether disclosing wage has a higher expected value than concealing due to the insufficient information on $P(W_E)$.

Job $V+1$ to job $T$ is the set where the wage $w$ is higher than all expected wages and falls in the productivity-dominant areas; in other words, both the expected callback probabilities and the proposed wages increase with wage. But it is not clear which job has the maximized expected value, as $P^j(w) > P^{j+1}(w)$ and $O^j(w) < O^{j+1}(w)$. In this range, disclosing is a dominant strategy: For every job $j$ $(V+1 \leq j \leq T)$, as $W^j_E < w$, we have $P^j(W^j_E) < P^j(w)$ and $O^j(W^j_E) < O^j(w)$, so $P^j(W^j_E)O^j(W^j_E) < P^j(w)O^j(w)$, implying that the expected value of disclosing is always higher than concealing, $U^D_{(V+1,T)} > U^C_{(V+1,T)}$.

For job $T+1$ to job $K$, the wage $w$ is in the productivity-dominant areas and is below all the expected wages. Similar to the second case, better jobs can propose a higher payment, $O^j(w) < O^{j+1}(w)$, but the probability of callback reduces with the increase of the firms' quality $P^j(w) > P^{j+1}(w)$, so we do not know which job has the highest expected value. However, opposite to the second case, concealing is the best response as it increases both the proposed wage and the chance of getting a callback: For every job $j$ $(T+1 \leq j \leq K)$, given $W^j_E > w$, $P^j(W^j_E) > P^j(w)$ and $O^j(W^j_E) > O^j(w)$, so $P^j(W^j_E)O^j(W^j_E) > P^j(w)O^j(w)$, implying that the expected value of concealing is always higher than disclosing, $U^D_{(T+1,K)} < U^C_{(T+1,K)}$.

We can obtain a simplified expression for the expected utility of disclosing:

$$U^D = U^D_{(v,T)} = max\{P^V(w)O^V(w), \ldots, P^T(w)O^T(w)\} \tag{B.12}$$

Figure B.3: Values of Disclosing and Concealing

Figure B.3a: Concealing Has a Higher Utility

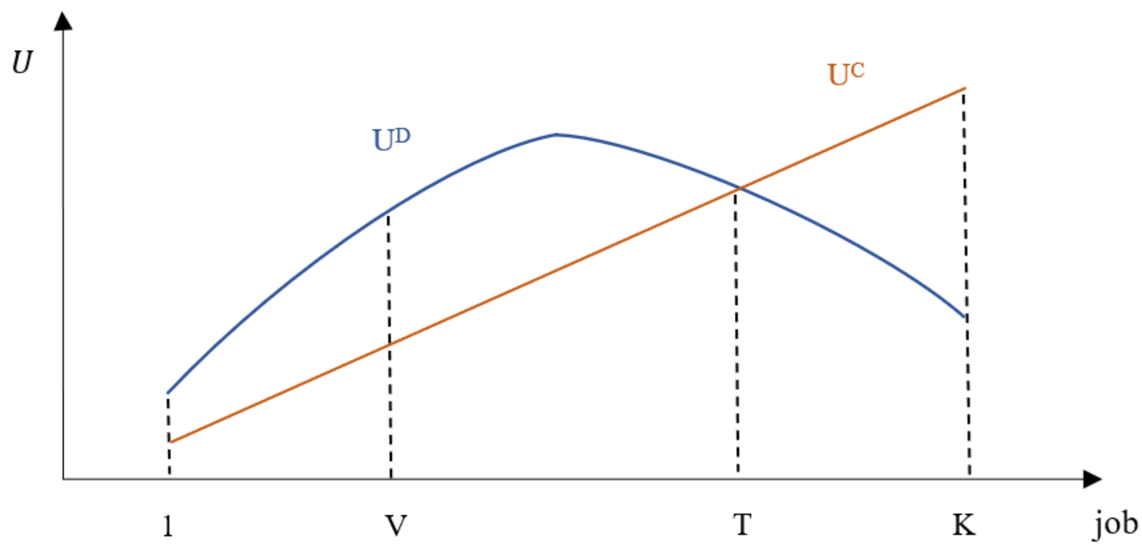



Figure B.3b: Disclosing Has a Higher Utility

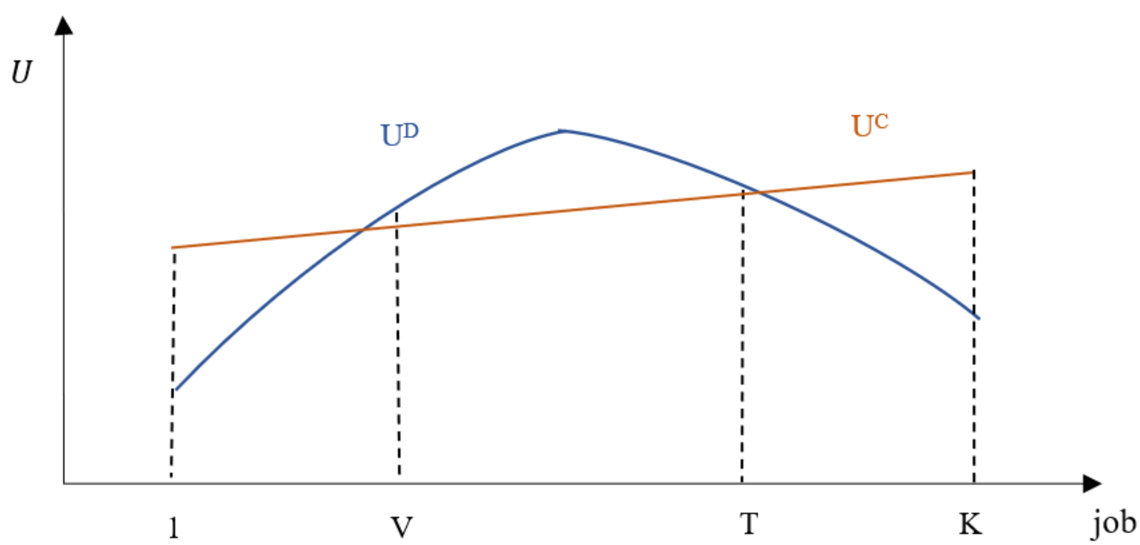


Figure B.3 shows an example of how the highest value of two options evolves with respect to the firm's quality index $j$. The expected utility of disclosing $U^D$ stems from jobs where the worker is better than the expected wage, but not over-qualified (in the outside option range). If $U^D$ is higher than the expected value of hiding wage $U^C$, the worker will choose to disclose his current wage. Put in another way, the more jobs in the range of $(V, T)$, the more likely the worker will disclose his wage. We can see this point from two extreme cases:

*Case 1:* $T = V = 0$, always concealer. When the worker's wage is lower than the smallest expected wage of jobs in his application set, $w < W_E^1 < \ldots < W_E^K$, he is an always concealer. For every job, $O(W_E) > O(w)$ due to the monotonicity of proposed wage, and $P(W_E) > P(w)$ because $w < W_E$ and $W_E$ is in the productivity-dominant area. Thus each element in $U^C$ will dominant the corresponding one in $U^D$, so concealing current wages is the best strategy for workers in this situation.

*Case 2:* $V = 0, T = K$, always discloser. When the worker's wage is higher than the largest expected wage and in the productivity-dominant areas in all the jobs in his application set, $W_E^1 < \ldots < W_E^K < w$, he is an always discloser. For every job, disclosing wage leads to a higher callback probability $P(W_E) < P(w)$ as well as a higher proposed wage $O(W_E) < O(w)$, then the worker will always disclose his wage information.

Wage disclosure can be driven by either a lower $V$ or a higher $T$. As mentioned above, $V$ is the number of jobs that the worker's wage is in their outside option-dominant areas, and a small $V$ indicates that there are few jobs that the worker is very overqualified in his application set. $T$ is the number of jobs with expected wages lower than the worker's actual wage, and a large $T$ implies that the worker has a relatively higher wage than the workers applied the same jobs as him. To conclude, given the same application set, compared to the low-wage workers whose wages are lower than the expectations of more jobs, high-wage workers (with higher $T$) are more likely to disclose their current wages to signal their better-than-expected productivity.[8]

## B.2 Descriptive Statistics

[8]In addition, workers who apply more aggressively (with small $V$), are more likely to disclose their current wages.

Table B.1: Descriptive Statistics: Applicant Sample

| | (1) Disclose | (2) Conceal | (3) All |
|---|---|---|---|
| **Disclose Current Wage** | **1.000** | **0.000** | **0.400** |
| Current Wage (annual, in 10,000 RMB) | 15.63 | 19.08 | 17.70 |
| **Demographics** | | | |
| Male | 0.671 | 0.610 | 0.634 |
| Age | 31.02 | 32.29 | 31.78 |
| Single | 0.170 | 0.154 | 0.161 |
| Married | 0.164 | 0.233 | 0.206 |
| Confidential Marital Status | 0.665 | 0.612 | 0.634 |
| **Employment Status** | | | |
| Employed, intensively search | 0.145 | 0.164 | 0.156 |
| Employed, moderately search | 0.550 | 0.568 | 0.561 |
| Employed, no plan to switch to new jobs | 0.000 | 0.001 | 0.001 |
| Unemployed | 0.305 | 0.268 | 0.282 |
| **Working History** | | | |
| Years of working experience | 8.275 | 9.126 | 8.786 |
| Tenure of the last job (years) | 2.758 | 2.610 | 2.669 |
| Tenure of the second last job (years) | 1.978 | 2.242 | 2.136 |
| **Education** | | | |
| Post doc | 0.000 | 0.001 | 0.001 |
| Phd | 0.004 | 0.006 | 0.005 |
| MBA/EMBA | 0.017 | 0.034 | 0.027 |
| Master | 0.112 | 0.194 | 0.161 |
| Bachelor | 0.604 | 0.613 | 0.610 |
| Tech collage | 0.233 | 0.141 | 0.178 |
| Secondary | 0.017 | 0.006 | 0.011 |
| High School | 0.012 | 0.004 | 0.008 |
| **Highest Degree** | | | |
| Domestic 1-19 | 0.060 | 0.091 | 0.079 |
| Domestic 20 - 39 | 0.052 | 0.069 | 0.063 |
| Domestic 40 - 59 | 0.032 | 0.039 | 0.036 |
| Domestic 60 - 99 | 0.066 | 0.088 | 0.079 |

| | | | |
|---|---|---|---|
| Domestic 100 - 199 | 0.105 | 0.115 | 0.111 |
| Domestic 200+ | 0.234 | 0.216 | 0.223 |
| World 1-19 | 0.000 | 0.001 | 0.001 |
| World 20 - 39 | 0.002 | 0.003 | 0.003 |
| World 40 - 59 | 0.006 | 0.012 | 0.009 |
| World 60 - 99 | 0.005 | 0.009 | 0.007 |
| World 100 - 199 | 0.018 | 0.029 | 0.025 |
| World 200+ | 0.116 | 0.155 | 0.140 |
| 985/211 Project University | 0.253 | 0.347 | 0.309 |
| Tongzhao degree | 0.834 | 0.834 | 0.834 |
| **Match** | | | |
| Province: current and desired | 0.869 | 0.854 | 0.860 |
| City: current and desired | 0.797 | 0.774 | 0.783 |
| Main industry: current and desired | 0.636 | 0.676 | 0.660 |
| Sub industry: current and desired | 0.323 | 0.359 | 0.344 |
| Main occupation: current and desired | 0.363 | 0.412 | 0.392 |
| Sub occupation: current and desired | 0.173 | 0.198 | 0.188 |
| Desired Wage (annual, in 10,000 RMB) | 17.15 | 23.83 | 21.16 |
| Disclose Desired Wage | 0.676 | 0.292 | 0.446 |
| **Website Classification** | | | |
| Resume Completeness Score | 630.5 | 644.0 | 638.6 |
| Elite Resume | 0.798 | 0.877 | 0.845 |
| Membership | 0.100 | 0.127 | 0.116 |
| Days from Registration | 551.9 | 819.2 | 712.3 |
| **Number of Applications** | 8.804 | 9.068 | 8.962 |
| | | | |
| **Sample Size** | 376,663 | 565,070 | 941,733 |

Note:

1. In addition to the variables listed above, we observe every applicant's current city and province, current main and sub- industries, current main and sub- occupations, desired working city and province, desired main and sub- industries, desired main and sub- occupations, and the sub-industries of the last two jobs in the working history.

2. Applicants can revise their profiles at any time, and the displayed statistics are from the resumes that applicants sent in their first job applications in our data

Table B.2: Descriptive Statistics: Job Ad Sample

| | Mean |
|---|---|
| **Posted Wage** | |
| Wage lower bound (annual, in 10,000 RMB) | 14.75 |
| Wage upper bound (annual, in 10,000 RMB) | 23.62 |
| Posted Wage Invisible to Worker | 0.240 |
| **Age Requirement** | |
| No requirement | 0.962 |
| Age lower bound | 24.03 |
| Age upper bound | 42.31 |
| **Experience Requirement** | |
| No requirement | 0.128 |
| Experience lower bound | 3.683 |
| **Education Requirement** | |
| No requirement | .065 |
| Post doc | .000 |
| Phd | .002 |
| MBA/EMBA | .000 |
| Master | .023 |
| Bachelor | .614 |
| Tech collage | .288 |
| Secondary | .009 |
| High School | .000 |
| Tongzhao Degree | .305 |
| **Gender Requirement** | |
| Male | .008 |
| Female | .002 |
| **Subordinate** | |
| No | .897 |
| Mean | 372 |
| **Feedback Days** | 4.30 |
| **Number of Applications** | 26.0 |
| | |
| **Sample Size** | 328,921 |

Note: In addition to the variables listed above, we can observe the main and sub- industry, main and sub- occupations, province and city, the open time, refresh time and close date for every job position.

Table B.3: Descriptive Statistics: Firm Sample

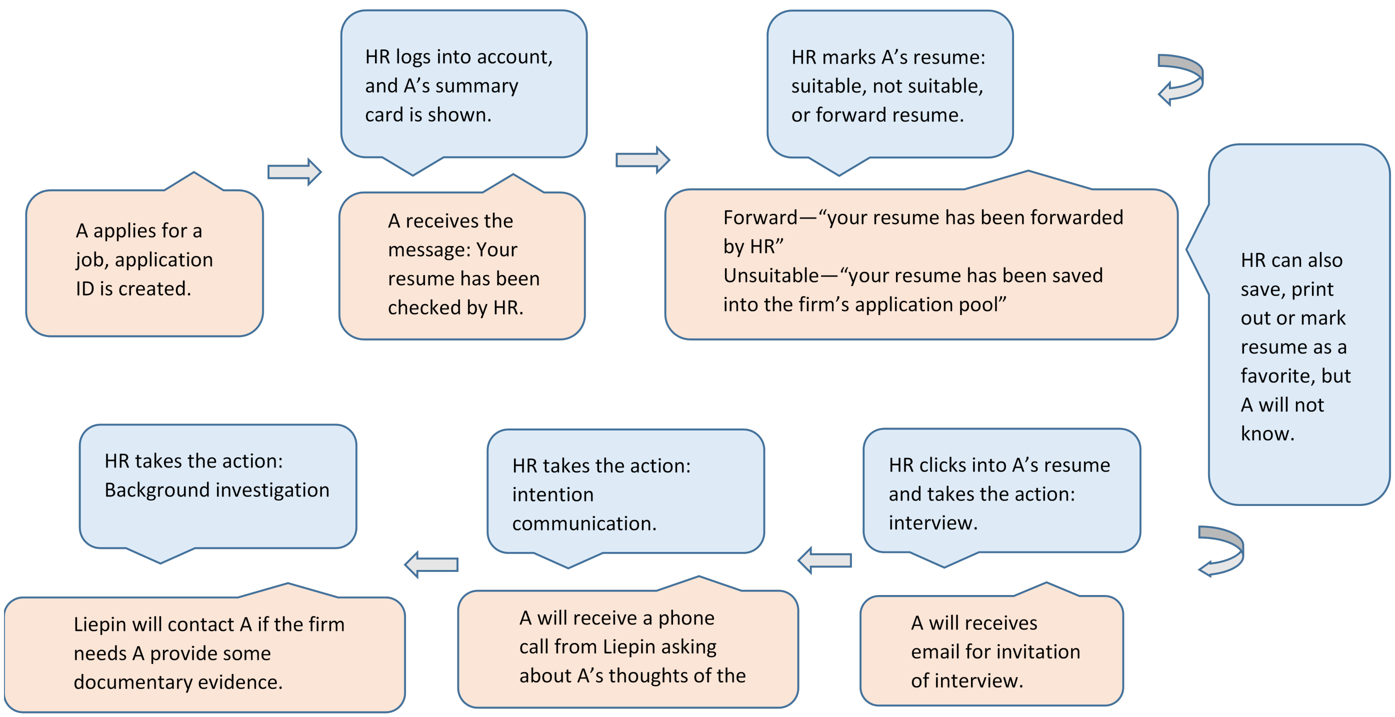
| | Mean |
|---|---|
| **Frim Size (Number of Employees)** | |
| 100-499 | 0.531 |
| 500-999 | 0.170 |
| 1000-2000 | 0.125 |
| 2000-5000 | 0.081 |
| 5000-10000 | 0.040 |
| 100000+ | 0.054 |
| **Firm Type** | |
| Foreign-funded | 0.120 |
| Sino-foreign joint | 0.073 |
| Private | 0.581 |
| State-owned | 0.086 |
| Public | 0.110 |
| Government/non-profit | 0.001 |
| Institution | 0.003 |
| Others | 0.027 |
| | |
| **Sample Size** | 19,264 |
| **Number of Hiring Agents** | 39,493 |

Table B.4: Descriptive Statistics: Viewed Application Sample

| | Percent |
|---|---|
| **Application Behaviors** | |
| Batch Apply | 0.274 |
| Use Lens Once | 0.032 |
| Use Lens Twice | 0.008 |
| **Application Results** | |
| Target | 0.048 |
| Save (Download) | 0.080 |
| Unsuitable | 0.432 |
| **Job Match Indicators** | |
| Over-educated | 0.321 |
| Under-educated | 0.087 |
| Over-age | 0.007 |
| Under-age | 0.003 |
| Mis-tongzhao degree | 0.062 |
| Under-experienced | 0.088 |
| Mis-gender | 0.003 |
| Current wage in the range of posted wage | 0.316 |
| Desired wage in the range of posted wage | 0.325 |
| | |
| City: job and current | 0.734 |
| City: job and desired | 0.747 |
| Province: job and current | 0.817 |
| Province: job and desired | 0.832 |
| Main industry: job and current | 0.428 |
| Main industry: job and desired | 0.561 |
| Sub industry: job and current | 0.298 |
| Sub industry: job and desired | 0.537 |
| Main Occupation: job and current | 0.464 |
| Main Occupation: job and desired | 0.527 |
| Sub Occupation: job and current | 0.218 |
| Sub Occupation: job and desired | 0.375 |
| | |
| **Sample Size** | 3,542,049 |

## B.3 Contexts in Liepin.com

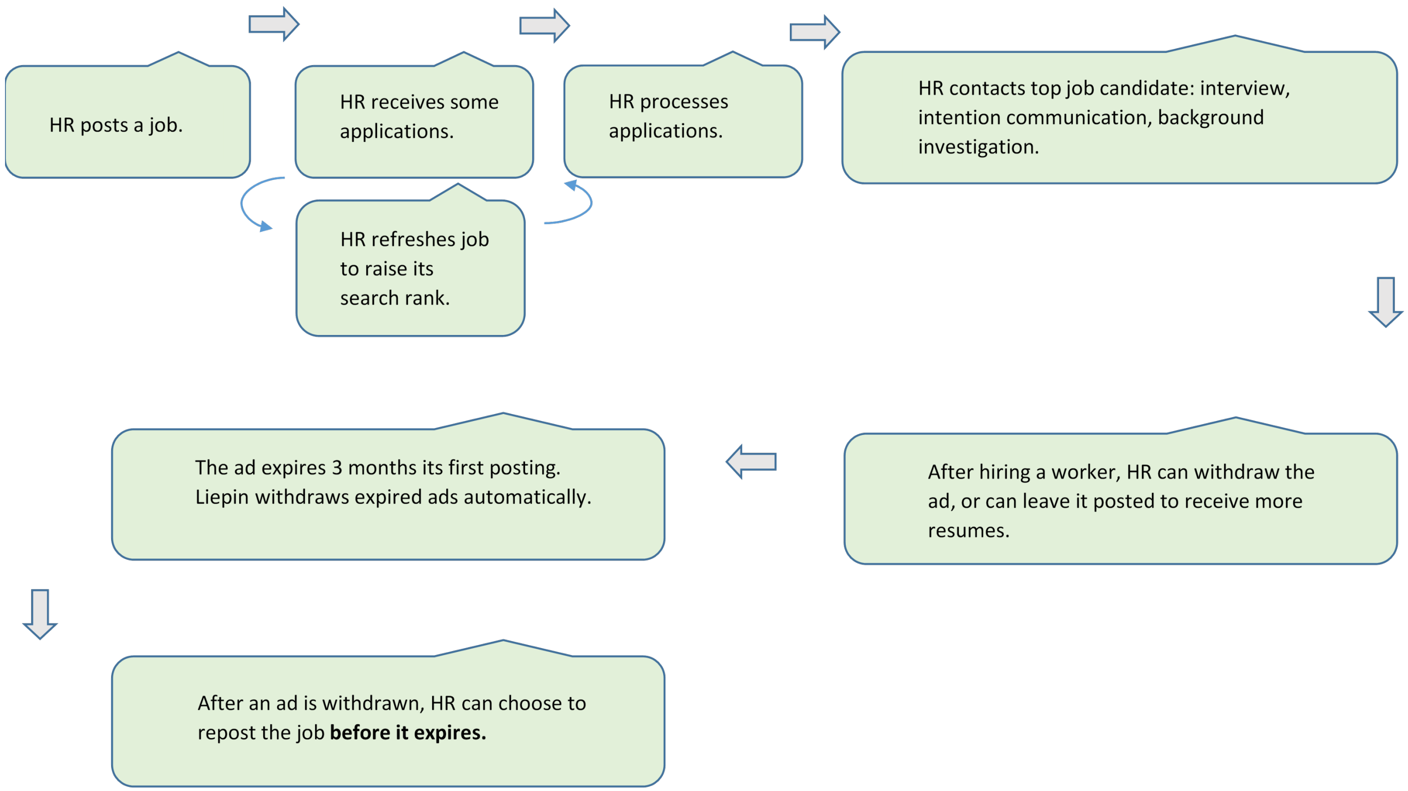
Figure B.4: Timeline of an Application Submitted in Liepin.com




Notes:

1. A denotes an applicant; HR denotes the hiring agent.

Figure B.5: Timeline of a Job Posted in Liepin.com



### B.3.1 Application Processing by Hiring Agents

When one application has been sent to the hiring agent, the applicant's resume will be saved into the hiring agent's account. However, not every application is guaranteed to be checked by hiring agents. The checking process can be divided into three stages. The first stage is that the applicant's resume is displayed as a summary card on the homepage of the hiring agent. The summary card only contains limited information including the applicant's photo, gender, education level, age, years of working experience, company name of the recent job, and location, but importantly, without any information of worker's wage. Summary cards are ranked by the application time, and the latest application is shown at the top. Each webpage displays 10 summary cards. If more new applications arrive, the application will be crowded out to the later pages, and the hiring agent may never see it, even as a summary card. When the applicant's summary card shows up on the hiring agent's homepage, he will receive a message such that "Your application has been received by HR". The proportion of applications that have displayed as summary cards on the hiring agent's page is 60.86%. This employer's response is almost random since the only factor considered into the ranking of summary cards is the application time.

The second stage is hiring agents choose to view the resumes. After seeing the summary card, the hiring agent can click into the summary card and view the whole resume of the applicant. Conditional on shown as summary cards, the proportion of applications viewed by hiring agents is 67.93%. This response should be affected only by the applicant's characteristics that are included in the summary card.

The third stage is hiring agents process the viewed applications. After viewing the full resume, hiring agents can process the application through the website by marking the applicant as target candidate, downloading the resume (saving as PDF version), or marking the applicant as unsuitable candidate. Conditional on viewed by hiring agent, 4.81% of applications are marked as target, 8.00% are saved as PDF, and 43.21% are marked as unsuitable.

To sum up, the employers' response variables recording the hiring agent's actions in different processing stages are summarized as following:

Figure B.6: Application Processing by Hiring Agents

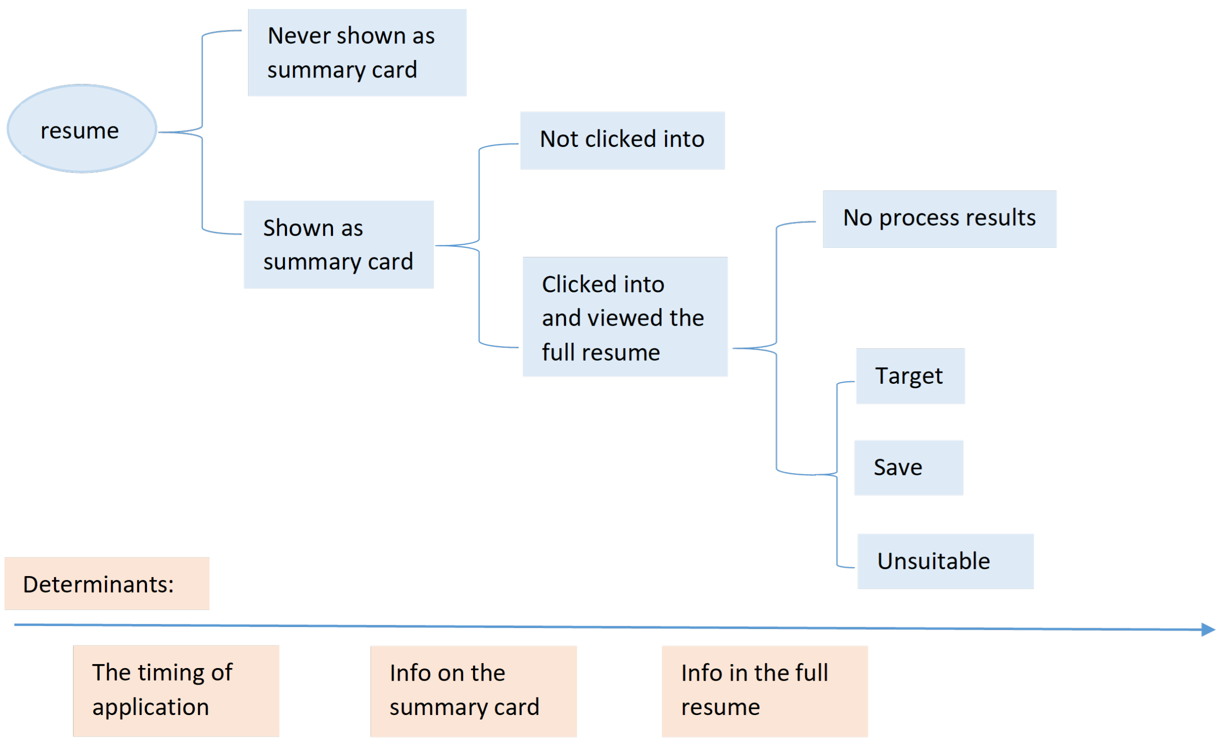


Table B.5: Descriptive Statistics: Full Application Sample

| | Percent |
|---|---|
| **Application Behaviors** | |
| Batch Apply | 0.298 |
| Use Lens Once | 0.026 |
| Use Lens Twice | 0.006 |
| **Application Results** | |
| Shown as Summary Card | 0.609 |
| Viewed by HR | 0.413 |
| Target | 0.020 |
| Save (Download) | 0.037 |
| Unsuitable | 0.181 |
| **Job Match Indicators** | |
| Over-educated | 0.304 |
| Under-educated | 0.090 |
| Over-age | 0.007 |
| Under-age | 0.002 |
| Mis-tongzhao degree | 0.061 |
| Under-experienced | 0.098 |
| Mis-gender | 0.004 |
| Current wage in the range of posted wage | 0.294 |
| Desired wage in the range of posted wage | 0.310 |
| | |
| City: job and current | 0.733 |
| City: job and desired | 0.750 |
| Province: job and current | 0.811 |
| Province: job and desired | 0.829 |
| Main industry: job and current | 0.413 |
| Main industry: job and desired | 0.540 |
| Sub industry: job and current | 0.286 |
| Sub industry: job and desired | 0.534 |
| Main Occupation: job and current | 0.453 |
| Main Occupation: job and desired | 0.498 |
| Sub Occupation: job and current | 0.205 |
| Sub Occupation: job and desired | 0.352 |
| | |
| **Sample Size** | 8,488,353 |

### B.3.2 Wage Disclosure and Viewed Applications

As noted above, since the wage information is not included in the summary card, worker's wage revealing should have no effect on the hiring agent's decision on whether to click into the summary card and view the full resume. We test this by regressing whether the application is viewed by the hiring agent on a set of job's characteristics and worker's information in summary card, plus the decision on wage disclosure in the sample of applications that have shown as summary cards in the hiring agent's account:

$$Y_i = \alpha_0 + \alpha_1 * DiscloseWage_i + \alpha_2 * X_i + FE + e_i \tag{B.13}$$

Table B.6 shows the regression results of the above specification, where the independent variable $Y_i$ is 1 if hiring agent has ever clicked into the summary card and viewed the full resume of application $i$. In column 1, we control for the applicant's wage level and his wage revealing choice. Column 2 controls for the applicant's gender and job's characteristics including requirements for age, education, and working experience; the offered wage range and whether the wage is visible to applicants, the number of position's subordinates and the reported feedback days. Column 3 controls for the match indicators between the applicant and the job: whether the applicant satisfies the job's gender, education, age and experience requirements, and whether the current and desired location (industry, occupation) are consistent with the location (industry, occupation) of the job. Column 4 adds variables for applicant's other characteristics in the summary card including the industry and tenure of the last two jobs, education quality, and controls for batch apply and the usage of job lens. Column 5 adds job's location, industry, occupation and firm fixed effects, and the fixed effect for the date of application. Column 6 replaces job's characteristics and location, industry, occupation and firm fixed effects with job fixed effect. In column 7, we drop worker's characteristics and include fixed effects for time, job and worker.

According to the estimates in Table B.6, worker's decision of wage disclosure has no effect on the probability of being viewed by hiring agents under any sets of controls, and the absolute value of wage does not significantly affect the view decision of hiring agents after controlling for job's characteristics and firm fixed effect. In general, employers prefer to click into the summary cards of male applicants than the same female summary cards, and the alignments between job's requirements and applicant's characteristics, including age, education, working experience, current working locations, industry and occupation will help the applicant win the first glance from hiring agents.

Table B.6: Effects of Wage Disclosure on the Probability of Resume Being Viewed by Hiring Agents

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| Disclose | -0.0004 | -0.0005 | -0.0003 | -0.0001 | 0.0005 | 0.0016 | 0.0009 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.002) | (0.001) |
| Wage | 0.0003*** | 0.0002*** | 0.0001** | 0.0000* | 0.0001 | 0.0001 | |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.001) | (0.001) | |
| Male | | 0.0047*** | 0.0055*** | 0.0017** | 0.0019*** | 0.0014*** | |
| | | (0.002) | (0.002) | (0.001) | (0.001) | (0.001) | |
| Worker over edu | | | 0.0052*** | 0.0076*** | 0.0071*** | 0.0001 | 0.0030*** |
| | | | (0.001) | (0.001) | (0.000) | (0.002) | (0.001) |
| Worker under edu | | | -0.0059*** | -0.0148*** | -0.0170*** | -0.0053** | -0.0109*** |
| | | | (0.002) | (0.001) | (0.001) | (0.002) | (0.001) |
| Worker over age | | | -0.0122 | -0.0109* | -0.0216*** | -0.0010 | -0.0103** |
| | | | (0.009) | (0.007) | (0.004) | (0.007) | (0.004) |
| Worker under age | | | -0.0255* | -0.0256*** | -0.0239*** | -0.0225*** | -0.0218*** |
| | | | (0.014) | (0.008) | (0.005) | (0.008) | (0.005) |
| Worker under exp | | | -0.0070*** | -0.0121*** | -0.0143*** | -0.0103*** | -0.0111*** |
| | | | (0.002) | (0.001) | (0.001) | (0.001) | (0.001) |
| Batch apply | | | -0.0302*** | -0.0203*** | -0.0139*** | -0.0146*** | -0.0074*** |
| | | | (0.001) | (0.001) | (0.000) | (0.001) | (0.001) |
| Lens*1 | | | 0.0330*** | 0.0344*** | 0.0239*** | 0.0394*** | 0.0256*** |
| | | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Lens*2 | | | 0.0189*** | 0.0195*** | 0.0117*** | 0.0264*** | 0.0157*** |
| | | | (0.003) | (0.002) | (0.002) | (0.003) | (0.002) |
| Job Characteristics | | Yes | Yes | Yes | Yes | | |
| Job Match | | | Yes | Yes | Yes | Yes | Yes |
| Worker Characteristics | | | | Yes | Yes | Yes | |
| City, Time, Ind & Occ FE | | | | | Yes | | |
| Firm FE | | | | | Yes | | |
| Job FE | | | | | | Yes | Yes |
| Worker FE | | | | | | | Yes |
| 'Effective' *N* | 5,166,319 | 5,166,319 | 5,166,319 | 5,165,774 | 5,114,629 | 4,883,726 | 4,825,933 |
| $R^2$ | 0.000 | 0.013 | 0.014 | 0.275 | 0.596 | 0.368 | 0.655 |

Standard errors in parentheses, clustered by job. *** p<0.01, ** p<0.05, * p<0.1

### B.3.3 Job Lens

The job lens is a function in Liepin.com, aiming at providing a rough evaluation of the worker's competitiveness and match degree between the worker and the job, and any registered job seekers can use it for free. In the job ad page, below the job requirement section, there is a button for lens with a notification "job lens: check the analysis of your competitiveness". After clicking that button, a small report will show up, and an example is presented in Figure B.7.

The first part of the report is position analysis. Position analysis includes two subsections, 1) activeness of job (low, median, high), with the last time when the hiring agent logged in, the last time when the hiring agent processed the job applications and how long it takes on average for the hiring agent to give feedback. 2) the popularity of job (low, median, high) based on the how the many applications that the job has received.

The second part is the competitiveness of job seeker. Based on the resume information, it lists the relative rank of the job seeker among all the received applicants for this job. First, it gives a total score of matching (low, median, high), and lists the requirements of the job, the satisfied requirements will be marked with $\surd$, the unsatisfied ones will have $\times$. Second, it reports the percentile of the jobseeker relative to the existing applicant pool: the distribution of applicants' working experience, the distribution of education level, the distribution of firms that the applicants are from (world top 500, China top 500, public company), the distribution of applicants' desired wage, the distribution of applicants' industry, and the distribution of their overseas working experience.

Figure B.7: An Example of Job Lens Report

## B.4 Wage Disclosure Behaviors

### B.4.1 Predicted Wage on Worker Level

In Section 2.5.1, we adopt worker's predicted wage as the cutoff for high-wage workers in indicator 3 (HighWage3). To reduce the prediction errors, we regress the worker's current wage on a set of his characteristics based on the sample of workers similar to him, in which the similar workers are defined as those who have ever applied the same jobs as him. More specifically, for each worker, we find all the jobs that the worker applied for, and collect all the applicants to these jobs, then the worker's predicted wage is estimated by the following regression in the sample of the applicants:

$$wage_j = \beta_0 + \beta_1 X_j + FE + e_j \tag{B.14}$$

$X_j$ is worker $j$'s demographics and characteristics including gender, quadratic in age, and interactions between age and gender, marital status and interactions between marital status and gender, quadratic in years of working experience, education level, school quality, employment status, tenure and industry in last two jobs, whether the worker is an elite, and the current-desired match variables. Fixed effects include worker's current location, industry and occupation.

We run this regression for every worker in the sample of the similar workers and predict worker's wage. Because this method requires a sufficient number of similar workers to run the regression, there are a small proportion of workers (5.7%) do not have the predicted wages, so HighWage3 is missing for these workers.

In Table B.7, we list the average effects and standard errors of some characteristics on worker's wage (i.e., the mean and the standard errors of the coefficients from 888,054 regressions).

Table B.7: Average Effects on Wage (Worker Level)

| Variables | Average Coefficients |
|---|---|
| Male | 1.2073*** |
| | (0.047) |
| Age | 0.2576*** |
| | (0.008) |
| Age*Male | 0.0794*** |
| | (0.006) |
| Single | 0.0843*** |
| | (0.021) |
| Married | 0.3387*** |
| | (0.076) |
| Single*Male | 0.0617*** |
| | (0.020) |
| Married*Male | 0.8518*** |
| | (0.096) |
| Experience | 1.0653*** |
| | (0.010) |
| Post doc | 11.2018*** |
| | (0.971) |
| Phd | 10.654*** |
| | (0.430) |
| MBA | 9.4871*** |
| | (0.266) |
| Master | 4.3785*** |
| | (0.177) |
| Bachelor | 2.3347*** |
| | (0.153) |
| Tech college | 0.3159** |
| | (0.139) |
| Secondary | 0.0439* |
| | (0.023) |
| Number of Workers | 888,054 |

Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

### B.4.2 Complete Version of Table 2.1

Table B.8 – B.10 presents the complete version of Table 2.2 in Panel A, B, and C, respectively. We find that younger, more experienced and less educated people are more likely to reveal their wages, and applicants who graduated from better universities are less likely to reveal their current wages. Unemployed workers, and workers who send out more job applications are more willing to reveal their current wages. With regard to the match variables, job seekers act differently between jumping to new fields and moving to other places. Job seekers who want to work in different locations are more likely to conceal their current wages, while people who want to work in different industries or occupations are more likely to reveal their current wages to future employers, and applicants who are seeking for a big wage growth in their next jobs, are less likely to disclose their current wages. According to the website classification variables, elite applicants, applicants who have been registered on Liepin for a longer time, and applicants with more complete profiles are unlikely to disclose their current wages.

Table B.8: The Effect of Applicant's Characteristics on Wage Disclosure (Panel A)

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| HighWage1 | 0.0472*** | 0.0494*** | 0.0593*** | 0.0533*** | 0.0477*** | 0.0514*** | 0.0510*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Wage | | -0.0048*** | -0.0034*** | -0.0018*** | -0.0009*** | -0.0004*** | -0.0004*** |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Male | | 0.0826*** | 0.0942*** | 0.0734*** | 0.0912*** | 0.0843*** | 0.0884*** |
| | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Age | | | -0.0093*** | -0.0130*** | -0.0103*** | -0.0110*** | -0.0106*** |
| | | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Age^2 | | | 0.0005*** | 0.0004*** | 0.0003*** | 0.0003*** | 0.0003*** |
| | | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Single | | | 0.0062*** | 0.0060*** | -0.0183*** | -0.0221*** | -0.0171*** |
| | | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Married | | | -0.0597*** | -0.0458*** | -0.0197*** | -0.0202*** | -0.0171*** |
| | | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Experience | | | | 0.0074*** | 0.0109*** | 0.0100*** | 0.0077*** |
| | | | | (0.000) | (0.000) | (0.000) | (0.000) |
| Experience^2 | | | | -0.0001*** | -0.0002*** | -0.0002*** | -0.0001*** |
| | | | | (0.000) | (0.000) | (0.000) | (0.000) |
| Post doc | | | | -0.1781*** | -0.1529*** | -0.1538*** | -0.1406*** |
| | | | | (0.021) | (0.021) | (0.021) | (0.021) |
| Phd | | | | -0.2190*** | -0.1924*** | -0.1956*** | -0.1802*** |
| | | | | (0.009) | (0.009) | (0.009) | (0.009) |
| MBA | | | | -0.2296*** | -0.1806*** | -0.1880*** | -0.1745*** |
| | | | | (0.006) | (0.006) | (0.006) | (0.007) |
| Master | | | | -0.2422*** | -0.2039*** | -0.2074*** | -0.1853*** |
| | | | | (0.006) | (0.006) | (0.006) | (0.006) |
| Bachelor | | | | -0.1690*** | -0.1319*** | -0.1384*** | -0.1231*** |
| | | | | (0.006) | (0.006) | (0.006) | (0.006) |
| Tech college | | | | -0.0787*** | -0.0527*** | -0.0605*** | -0.0553*** |
| | | | | (0.006) | (0.006) | (0.006) | (0.006) |
| Secondary | | | | 0.0005 | 0.0045 | -0.0017 | -0.0007 |
| | | | | (0.007) | (0.007) | (0.007) | (0.007) |
| Tongzhao | | | | -0.0253*** | -0.0223*** | -0.0219*** | -0.0164*** |
| | | | | (0.001) | (0.001) | (0.001) | (0.001) |
| 985/211 | | | | -0.0119*** | -0.0114*** | -0.0141*** | -0.0117*** |
| | | | | (0.002) | (0.002) | (0.002) | (0.002) |
| Intensively Search | | | | -0.0210*** | 0.0005 | 0.0019 | 0.0041*** |
| | | | | (0.002) | (0.002) | (0.002) | (0.002) |
| Moderately Search | | | | -0.1531*** | -0.1246*** | -0.1218*** | -0.1111*** |
| | | | | (0.020) | (0.020) | (0.020) | (0.019) |
| Stay | | | | -0.0062*** | 0.0012 | -0.0009 | -0.0010 |
| | | | | (0.001) | (0.001) | (0.001) | (0.001) |
| Tenure 1 | | | | 0.0025*** | 0.0009*** | 0.0009*** | 0.0010*** |

| | | | | |
|---|---|---|---|---|
| | (0.000) | (0.000) | (0.000) | (0.000) |
| Tenure 2 | 0.0044*** | 0.0021*** | 0.0017*** | 0.0016*** |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| Domes 1-19 | -0.0247*** | -0.0188*** | -0.0146*** | -0.0125*** |
| | (0.003) | (0.003) | (0.003) | (0.003) |
| Domes 20-39 | -0.0310*** | -0.0263*** | -0.0205*** | -0.0154*** |
| | (0.003) | (0.003) | (0.003) | (0.003) |
| Domes 40-59 | -0.0188*** | -0.0144*** | -0.0107*** | -0.0107*** |
| | (0.003) | (0.003) | (0.003) | (0.003) |
| Domes 60-99 | -0.0360*** | -0.0315*** | -0.0205*** | -0.0186*** |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| Dome 100-199 | -0.0191*** | -0.0144*** | -0.0162*** | -0.0139*** |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| World 1-19 | -0.0981*** | -0.1101*** | -0.0952*** | -0.0849*** |
| | (0.017) | (0.017) | (0.017) | (0.017) |
| World 20-39 | -0.0317*** | -0.0405*** | -0.0411*** | -0.0367*** |
| | (0.010) | (0.010) | (0.010) | (0.010) |
| World 40-59 | -0.0262*** | -0.0301*** | -0.0274*** | -0.0269*** |
| | (0.005) | (0.005) | (0.005) | (0.005) |
| World 60-99 | -0.0337*** | -0.0350*** | -0.0270*** | -0.0261*** |
| | (0.006) | (0.006) | (0.006) | (0.006) |
| World 100-199 | -0.0266*** | -0.0279*** | -0.0150*** | -0.0123*** |
| | (0.004) | (0.004) | (0.004) | (0.004) |
| Elite | | -0.0097*** | -0.0097*** | 0.0066*** |
| | | (0.002) | (0.002) | (0.002) |
| Account days | | -0.9404*** | -0.9110*** | -0.9221*** |
| | | (0.008) | (0.008) | (0.008) |
| Resume Completeness | | -2.7647*** | -2.5389*** | -2.4203*** |
| | | (0.093) | (0.092) | (0.093) |
| Membership | | -0.0199*** | -0.0183*** | -0.0185*** |
| | | (0.002) | (0.002) | (0.002) |
| Number of Applications | | 0.0004*** | 0.0005*** | 0.0004*** |
| | | (0.000) | (0.000) | (0.000) |
| Log (wage gap) | | -0.0129*** | -0.0119*** | -0.0109*** |
| | | (0.000) | (0.000) | (0.000) |
| City Match | | 0.0289*** | 0.0399*** | 0.0409*** |
| | | (0.002) | (0.002) | (0.002) |
| Province Match | | -0.0042* | 0.0078*** | 0.0045* |
| | | (0.002) | (0.002) | (0.002) |
| Main Industry Match | | -0.0223*** | -0.0202*** | -0.0120*** |
| | | (0.002) | (0.002) | (0.002) |
| Sub-Industry Match | | -0.0161*** | -0.0191*** | -0.0223*** |
| | | (0.001) | (0.001) | (0.001) |
| Main Occupation Match | | -0.0233*** | -0.0225*** | -0.0198*** |
| | | (0.002) | (0.002) | (0.002) |
| Sub-Occupation Match | | -0.0184*** | -0.0180*** | -0.0115*** |
| | | (0.002) | (0.002) | (0.002) |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 'Effective' *N* | 941,733 | 941,733 | 941,733 | 941,733 | 941,733 | 941,697 | 941,695 |
| $R^2$ | 0.000 | 0.021 | 0.032 | 0.059 | 0.083 | 0.091 | 0.102 |

Standard errors in parentheses, clustered by worker's sub-occupation. *** p<0.01, ** p<0.05, * p<0.1

Note: In Table B.8-10, column 6 controls for the worker's location fixed effect, and column 7 further controls for the worker's industry and occupation fixed effects.

Table B.9: The Effect of Applicant's Characteristics on Wage Disclosure (Panel B)

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| HighWage2 | 0.0288*** | 0.0311*** | 0.0400*** | 0.0421*** | 0.0405*** | 0.0417*** | 0.0450*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Wage | | -0.0046*** | -0.0033*** | -0.0018*** | -0.0009*** | -0.0004*** | -0.0004*** |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Male | | 0.0822*** | 0.0935*** | 0.0726*** | 0.0907*** | 0.0839*** | 0.0880*** |
| | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Age | | | -0.0091*** | -0.0129*** | -0.0102*** | -0.0108*** | -0.0105*** |
| | | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Age^2 | | | 0.0005*** | 0.0004*** | 0.0003*** | 0.0003*** | 0.0003*** |
| | | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Single | | | 0.0045*** | 0.0047*** | -0.0195*** | -0.0233*** | -0.0182*** |
| | | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Married | | | -0.0596*** | -0.0459*** | -0.0196*** | -0.0201*** | -0.0171*** |
| | | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Experience | | | | 0.0075*** | 0.0110*** | 0.0101*** | 0.0078*** |
| | | | | (0.000) | (0.000) | (0.000) | (0.000) |
| Experience^2 | | | | -0.0001*** | -0.0002*** | -0.0002*** | -0.0001*** |
| | | | | (0.000) | (0.000) | (0.000) | (0.000) |
| Post doc | | | | -0.1862*** | -0.1597*** | -0.1611*** | -0.1460*** |
| | | | | (0.021) | (0.021) | (0.021) | (0.021) |
| Phd | | | | -0.2271*** | -0.1992*** | -0.2029*** | -0.1855*** |
| | | | | (0.009) | (0.009) | (0.009) | (0.009) |
| MBA | | | | -0.2366*** | -0.1865*** | -0.1942*** | -0.1789*** |
| | | | | (0.006) | (0.006) | (0.006) | (0.007) |
| Master | | | | -0.2488*** | -0.2098*** | -0.2135*** | -0.1897*** |
| | | | | (0.006) | (0.006) | (0.006) | (0.006) |
| Bachelor | | | | -0.1747*** | -0.1369*** | -0.1435*** | -0.1267*** |
| | | | | (0.006) | (0.006) | (0.006) | (0.006) |
| Tech college | | | | -0.0818*** | -0.0556*** | -0.0632*** | -0.0568*** |
| | | | | (0.006) | (0.006) | (0.006) | (0.006) |
| Secondary | | | | -0.0001 | 0.0038 | -0.0022 | -0.0010 |
| | | | | (0.007) | (0.007) | (0.007) | (0.007) |
| Tongzhao | | | | -0.0261*** | -0.0229*** | -0.0226*** | -0.0171*** |
| | | | | (0.001) | (0.001) | (0.001) | (0.001) |
| 985/211 | | | | -0.0119*** | -0.0114*** | -0.0140*** | -0.0118*** |
| | | | | (0.002) | (0.002) | (0.002) | (0.002) |
| Intensively Search | | | | -0.0220*** | -0.0003 | 0.0010 | 0.0032** |
| | | | | (0.002) | (0.002) | (0.002) | (0.002) |
| Moderately Search | | | | -0.1540*** | -0.1251*** | -0.1225*** | -0.1116*** |
| | | | | (0.020) | (0.020) | (0.020) | (0.019) |
| Stay | | | | -0.0079*** | -0.0001 | -0.0024** | -0.0026** |
| | | | | (0.001) | (0.001) | (0.001) | (0.001) |
| Tenure 1 | | | | 0.0024*** | 0.0009*** | 0.0008*** | 0.0009*** |

| | | | | |
|---|---|---|---|---|
| | (0.000) | (0.000) | (0.000) | (0.000) |
| Tenure 2 | 0.0044*** | 0.0022*** | 0.0017*** | 0.0016*** |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| Domes 1-19 | -0.0250*** | -0.0190*** | -0.0148*** | -0.0126*** |
| | (0.003) | (0.003) | (0.003) | (0.003) |
| Domes 20-39 | -0.0312*** | -0.0265*** | -0.0207*** | -0.0156*** |
| | (0.003) | (0.003) | (0.003) | (0.003) |
| Domes 40-59 | -0.0191*** | -0.0146*** | -0.0110*** | -0.0109*** |
| | (0.003) | (0.003) | (0.003) | (0.003) |
| Domes 60-99 | -0.0362*** | -0.0317*** | -0.0209*** | -0.0188*** |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| Dome 100-199 | -0.0193*** | -0.0147*** | -0.0164*** | -0.0140*** |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| World 1-19 | -0.0988*** | -0.1107*** | -0.0962*** | -0.0860*** |
| | (0.017) | (0.017) | (0.017) | (0.017) |
| World 20-39 | -0.0320*** | -0.0409*** | -0.0410*** | -0.0366*** |
| | (0.010) | (0.010) | (0.010) | (0.010) |
| World 40-59 | -0.0272*** | -0.0309*** | -0.0280*** | -0.0275*** |
| | (0.005) | (0.005) | (0.005) | (0.005) |
| World 60-99 | -0.0330*** | -0.0343*** | -0.0271*** | -0.0262*** |
| | (0.006) | (0.006) | (0.006) | (0.006) |
| World 100-199 | -0.0261*** | -0.0275*** | -0.0151*** | -0.0124*** |
| | (0.004) | (0.004) | (0.004) | (0.004) |
| Elite | | -0.0094*** | -0.0091*** | 0.0069*** |
| | | (0.002) | (0.002) | (0.002) |
| Account days | | -0.9434*** | -0.9143*** | -0.9257*** |
| | | (0.008) | (0.008) | (0.008) |
| Resume Completeness | | -2.7630*** | -2.5383*** | -2.4263*** |
| | | (0.093) | (0.093) | (0.093) |
| Membership | | -0.0200*** | -0.0184*** | -0.0186*** |
| | | (0.002) | (0.002) | (0.002) |
| Number of Applications | | 0.0004*** | 0.0005*** | 0.0005*** |
| | | (0.000) | (0.000) | (0.000) |
| Log (wage gap) | | -0.0132*** | -0.0124*** | -0.0112*** |
| | | (0.000) | (0.000) | (0.000) |
| City Match | | 0.0288*** | 0.0401*** | 0.0412*** |
| | | (0.002) | (0.002) | (0.002) |
| Province Match | | -0.0038* | 0.0079*** | 0.0046** |
| | | (0.002) | (0.002) | (0.002) |
| Main Industry Match | | -0.0220*** | -0.0199*** | -0.0114*** |
| | | (0.002) | (0.002) | (0.002) |
| Sub-Industry Match | | -0.0167*** | -0.0196*** | -0.0227*** |
| | | (0.001) | (0.001) | (0.001) |
| Main Occupation Match | | -0.0236*** | -0.0228*** | -0.0195*** |
| | | (0.002) | (0.002) | (0.002) |
| Sub-Occupation Match | | -0.0183*** | -0.0178*** | -0.0113*** |
| | | (0.002) | (0.002) | (0.002) |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 'Effective' *N* | 941,733 | 941,733 | 941,733 | 941,733 | 941,733 | 941,697 | 941,695 |
| $R^2$ | 0.001 | 0.020 | 0.031 | 0.058 | 0.082 | 0.090 | 0.101 |

Standard errors in parentheses, clustered by worker's sub-occupation. *** p<0.01, ** p<0.05, * p<0.1

Table B.10: The Effect of Applicant's Characteristics on Wage Disclosure (Panel C)

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| HighWage3 | 0.0292*** | 0.0870*** | 0.0645*** | 0.0310*** | 0.0314*** | 0.0154*** | 0.0157*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Wage | | -0.0052*** | -0.0038*** | -0.0018*** | -0.0009*** | -0.0001* | -0.0001 |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Male | | 0.0880*** | 0.0949*** | 0.0736*** | 0.0909*** | 0.0837*** | 0.0877*** |
| | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Age | | | -0.0067*** | -0.0121*** | -0.0096*** | -0.0104*** | -0.0100*** |
| | | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Age^2 | | | 0.0004*** | 0.0004*** | 0.0003*** | 0.0003*** | 0.0003*** |
| | | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Single | | | 0.0044*** | 0.0046*** | -0.0194*** | -0.0230*** | -0.0180*** |
| | | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Married | | | -0.0577*** | -0.0449*** | -0.0184*** | -0.0190*** | -0.0158*** |
| | | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Experience | | | | 0.0085*** | 0.0120*** | 0.0107*** | 0.0084*** |
| | | | | (0.000) | (0.000) | (0.000) | (0.000) |
| Experience^2 | | | | -0.0002*** | -0.0002*** | -0.0002*** | -0.0002*** |
| | | | | (0.000) | (0.000) | (0.000) | (0.000) |
| Post doc | | | | -0.1772*** | -0.1560*** | -0.1630*** | -0.1480*** |
| | | | | (0.021) | (0.021) | (0.021) | (0.021) |
| Phd | | | | -0.2174*** | -0.1946*** | -0.2044*** | -0.1872*** |
| | | | | (0.009) | (0.009) | (0.009) | (0.009) |
| MBA | | | | -0.2286*** | -0.1833*** | -0.1962*** | -0.1816*** |
| | | | | (0.007) | (0.006) | (0.006) | (0.007) |
| Master | | | | -0.2403*** | -0.2073*** | -0.2148*** | -0.1916*** |
| | | | | (0.006) | (0.006) | (0.006) | (0.006) |
| Bachelor | | | | -0.1701*** | -0.1379*** | -0.1460*** | -0.1300*** |
| | | | | (0.006) | (0.006) | (0.006) | (0.006) |
| Tech college | | | | -0.0815*** | -0.0596*** | -0.0665*** | -0.0610*** |
| | | | | (0.006) | (0.006) | (0.006) | (0.006) |
| Secondary | | | | 0.0001 | 0.0038 | -0.0018 | -0.0009 |
| | | | | (0.007) | (0.007) | (0.007) | (0.007) |
| Tongzhao | | | | -0.0238*** | -0.0213*** | -0.0219*** | -0.0164*** |
| | | | | (0.001) | (0.001) | (0.001) | (0.001) |
| 985/211 | | | | -0.0096*** | -0.0094*** | -0.0130*** | -0.0107*** |
| | | | | (0.002) | (0.002) | (0.002) | (0.002) |
| Intensively Search | | | | -0.0219*** | -0.0008 | 0.0003 | 0.0026* |
| | | | | (0.002) | (0.002) | (0.002) | (0.002) |
| Moderately Search | | | | -0.1476*** | -0.1191*** | -0.1197*** | -0.1091*** |
| | | | | (0.020) | (0.020) | (0.020) | (0.019) |
| Stay | | | | -0.0062*** | 0.0014 | -0.0021* | -0.0021* |
| | | | | (0.001) | (0.001) | (0.001) | (0.001) |
| Tenure 1 | | | | 0.0021*** | 0.0005** | 0.0006*** | 0.0007*** |

| | | | | |
|---|---|---|---|---|
| | (0.000) | (0.000) | (0.000) | (0.000) |
| Tenure 2 | 0.0042*** | 0.0019*** | 0.0017*** | 0.0016*** |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| Domes 1-19 | -0.0246*** | -0.0185*** | -0.0142*** | -0.0121*** |
| | (0.003) | (0.003) | (0.003) | (0.003) |
| Domes 20-39 | -0.0318*** | -0.0271*** | -0.0207*** | -0.0155*** |
| | (0.003) | (0.003) | (0.003) | (0.003) |
| Domes 40-59 | -0.0207*** | -0.0164*** | -0.0117*** | -0.0117*** |
| | (0.003) | (0.003) | (0.003) | (0.003) |
| Domes 60-99 | -0.0365*** | -0.0321*** | -0.0212*** | -0.0192*** |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| Dome 100-199 | -0.0201*** | -0.0157*** | -0.0167*** | -0.0143*** |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| World 1-19 | -0.0903*** | -0.1032*** | -0.0937*** | -0.0832*** |
| | (0.017) | (0.017) | (0.017) | (0.017) |
| World 20-39 | -0.0308*** | -0.0404*** | -0.0406*** | -0.0362*** |
| | (0.010) | (0.010) | (0.010) | (0.010) |
| World 40-59 | -0.0243*** | -0.0278*** | -0.0278*** | -0.0271*** |
| | (0.005) | (0.005) | (0.005) | (0.005) |
| World 60-99 | -0.0298*** | -0.0311*** | -0.0268*** | -0.0257*** |
| | (0.006) | (0.006) | (0.006) | (0.006) |
| World 100-199 | -0.0250*** | -0.0265*** | -0.0151*** | -0.0123*** |
| | (0.004) | (0.004) | (0.004) | (0.004) |
| Elite | | 0.0018 | 0.0004 | 0.0167*** |
| | | (0.002) | (0.002) | (0.002) |
| Account days | | -0.9496*** | -0.9196*** | -0.9319*** |
| | | (0.008) | (0.008) | (0.009) |
| Resume Completeness | | -2.7611*** | -2.5160*** | -2.4018*** |
| | | (0.093) | (0.093) | (0.093) |
| Membership | | -0.0200*** | -0.0184*** | -0.0186*** |
| | | (0.002) | (0.002) | (0.002) |
| Number of Applications | | 0.0004*** | 0.0005*** | 0.0005*** |
| | | (0.000) | (0.000) | (0.000) |
| Log (wage gap) | | -0.0135*** | -0.0127*** | -0.0116*** |
| | | (0.000) | (0.000) | (0.000) |
| City Match | | 0.0308*** | 0.0393*** | 0.0402*** |
| | | (0.002) | (0.002) | (0.002) |
| Province Match | | -0.0076*** | 0.0067*** | 0.0034 |
| | | (0.002) | (0.002) | (0.002) |
| Main Industry Match | | -0.0204*** | -0.0191*** | -0.0113*** |
| | | (0.002) | (0.002) | (0.002) |
| Sub-Industry Match | | -0.0165*** | -0.0192*** | -0.0226*** |
| | | (0.001) | (0.001) | (0.001) |
| Main Occupation Match | | -0.0227*** | -0.0225*** | -0.0193*** |
| | | (0.002) | (0.002) | (0.002) |
| Sub-Occupation Match | | -0.0198*** | -0.0185*** | -0.0121*** |
| | | (0.002) | (0.002) | (0.002) |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 'Effective' $N$ | 888,054 | 888,054 | 888,054 | 888,054 | 888,054 | 888,018 | 887,980 |
| $R^2$ | 0.001 | 0.026 | 0.033 | 0.058 | 0.082 | 0.089 | 0.100 |

Standard errors in parentheses, clustered by worker's sub-occupation. *** p<0.01, ** p<0.05, * p<0.1

### B.4.3 Heterogeneity of Wage Disclosure

Table B.11: The Effect of Gender on Wage Disclosure Decision

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| **Panel A** | | | | | | | |
| HighWage1*Male | -0.0231*** | -0.0100*** | -0.0147*** | -0.0116*** | -0.0115*** | -0.0108*** | -0.0121*** |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| HighWage1 | 0.0071*** | 0.0557*** | 0.0685*** | 0.0606*** | 0.0550*** | 0.0582*** | 0.0587*** |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| Male | 0.0744*** | 0.0871*** | 0.1007*** | 0.0786*** | 0.0963*** | 0.0892*** | 0.0939*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| **Panel B** | | | | | | | |
| HighWage2*Male | -0.0257*** | -0.0094*** | -0.0158*** | -0.0113*** | -0.0114*** | -0.0115*** | -0.0108*** |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| HighWage2 | -0.0199*** | 0.0370*** | 0.0498*** | 0.0492*** | 0.0477*** | 0.0489*** | 0.0519*** |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| Male | 0.0765*** | 0.0862*** | 0.1002*** | 0.0774*** | 0.0955*** | 0.0888*** | 0.0926*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| **Panel C** | | | | | | | |
| HighWage3*Male | -0.0399*** | -0.0120*** | -0.0147*** | -0.0113*** | -0.0107*** | -0.0115*** | -0.0104*** |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| HighWage3 | 0.0558*** | 0.0943*** | 0.0734*** | 0.0379*** | 0.0379*** | 0.0224*** | 0.0220*** |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| Male | 0.0818*** | 0.0932*** | 0.1012*** | 0.0784*** | 0.0955*** | 0.0886*** | 0.0921*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Demographics | | | Yes | Yes | Yes | Yes | Yes |
| Edu & Exp | | | | Yes | Yes | Yes | Yes |
| Classification & Match | | | | | Yes | Yes | Yes |
| Location FE | | | | | | Yes | Yes |
| Industry FE | | | | | | | Yes |
| Occupation FE | | | | | | | Yes |
| ‘Effective’ *N* | 941,733 | 941,733 | 941,733 | 941,733 | 941,733 | 941,697 | 941,695 |

Standard errors in parentheses, clustered by worker’s sub-occupation. *** p<0.01, ** p<0.05, * p<0.1

Note: Column 2 controls for worker’s current wage. Column 3-7 have the same specifications as Table 1.

Table B.12: The Effect of Unemployment on Wage Disclosure Decision

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| **Panel A** | | | | | | | |
| HighWage1* Unemp | 0.0168*** | 0.0115*** | 0.0230*** | 0.0177*** | 0.0169*** | 0.0172*** | 0.0144*** |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| HighWage1 | -0.0097*** | 0.0450*** | 0.0518*** | 0.0481*** | 0.0430*** | 0.0466*** | 0.0470*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Unemp | 0.0362*** | 0.0250*** | 0.0104*** | 0.0017 | 0.0087*** | 0.0075*** | 0.0066*** |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.001) | (0.001) | (0.001) |
| **Panel B** | | | | | | | |
| HighWage2* Unemp | 0.0106*** | 0.0024 | 0.0116*** | 0.0071*** | 0.0074*** | 0.0078*** | 0.0072*** |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| HighWage2 | -0.0344*** | 0.0302*** | 0.0363*** | 0.0400*** | 0.0385*** | 0.0395*** | 0.0430*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Unemp | 0.0376*** | 0.0309*** | 0.0182*** | 0.0083*** | 0.0029** | 0.0015 | 0.0016 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| **Panel C** | | | | | | | |
| HighWage3* Unemp | 0.0341*** | 0.0062*** | 0.0050** | 0.0047** | 0.0051** | 0.0091*** | 0.0095*** |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| HighWage3 | 0.0175*** | 0.0828*** | 0.0626*** | 0.0318*** | 0.0329*** | 0.0183*** | 0.0188*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Unemp | 0.0270*** | 0.0207*** | 0.0181*** | 0.0120*** | -0.0014 | 0.0057*** | 0.0053*** |
| | (0.002) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Demographics | | | Yes | Yes | Yes | Yes | Yes |
| Edu & Exp | | | | Yes | Yes | Yes | Yes |
| Classification & Match | | | | | Yes | Yes | Yes |
| Location FE | | | | | | Yes | Yes |
| Industry FE | | | | | | | Yes |
| Occupation FE | | | | | | | Yes |
| 'Effective' *N* | 941,733 | 941,733 | 941,733 | 941,733 | 941,733 | 941,697 | 941,695 |

Standard errors in parentheses, clustered by worker's sub-occupation. *** p<0.01, ** p<0.05, * p<0.1

Note: Column 2 controls for worker's gender and current wage. Column 3-7 have the same specifications as Table 1.

### B.4.4 Robustness Check on Worker's Wage Disclosure Decision

Figure B.8: Effects of HighWage by Industry

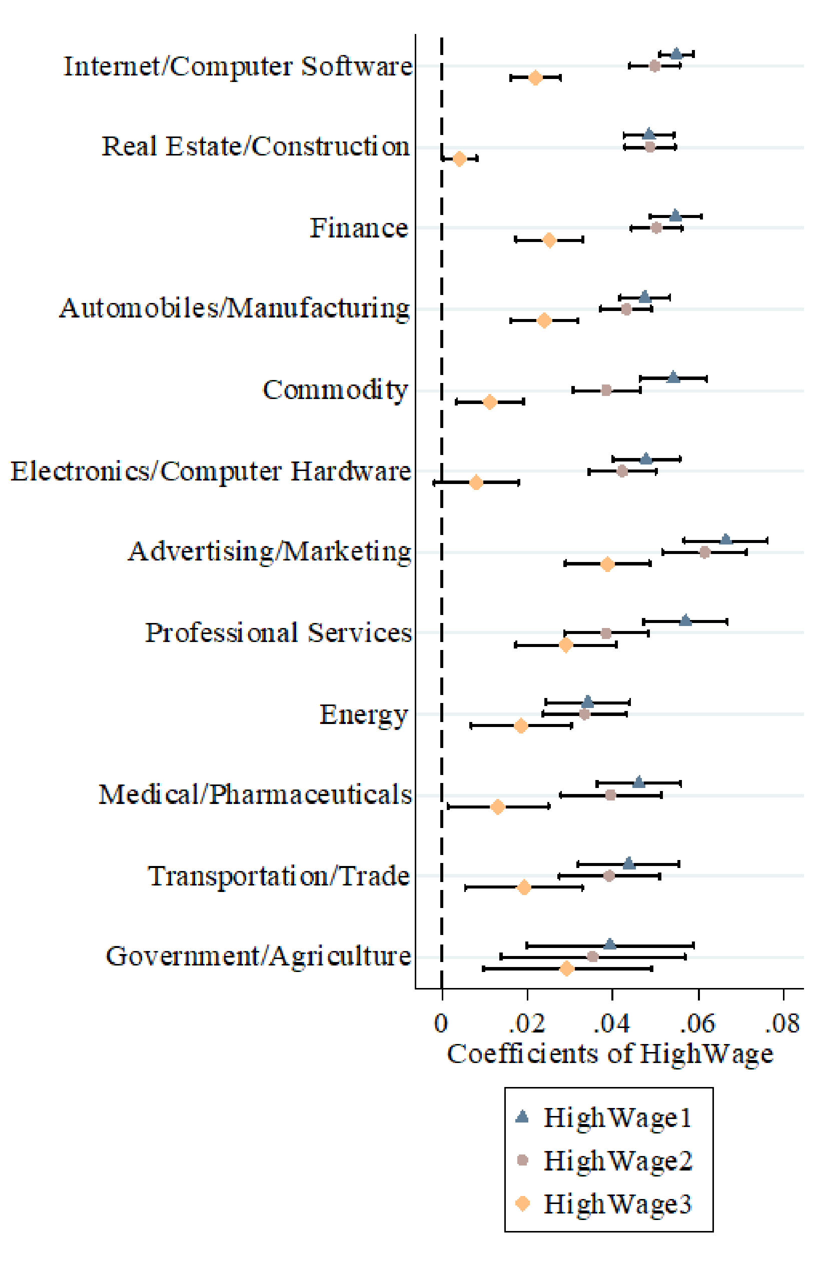


Figure B.9: Effects of HighWage by Occupation

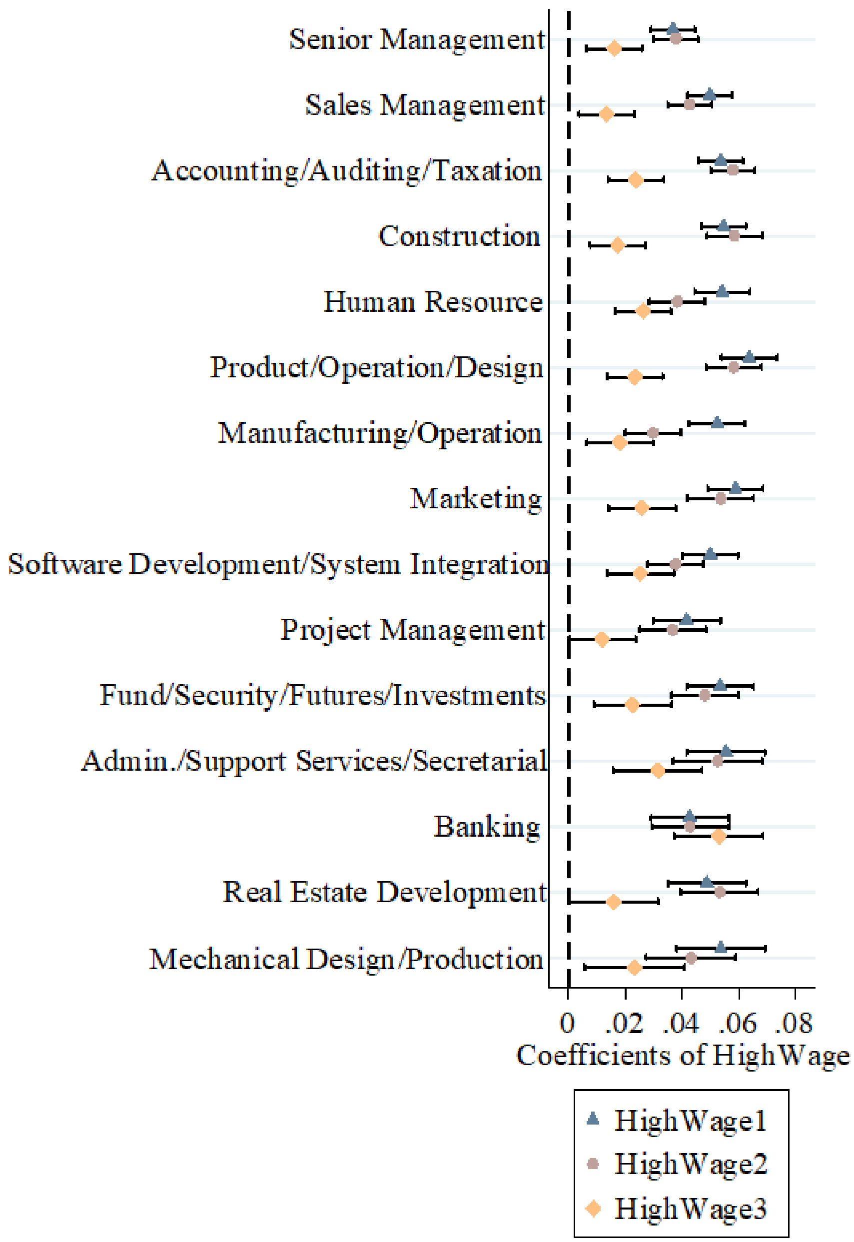


Figure B.10: Effects of HighWage by Wage Deciles

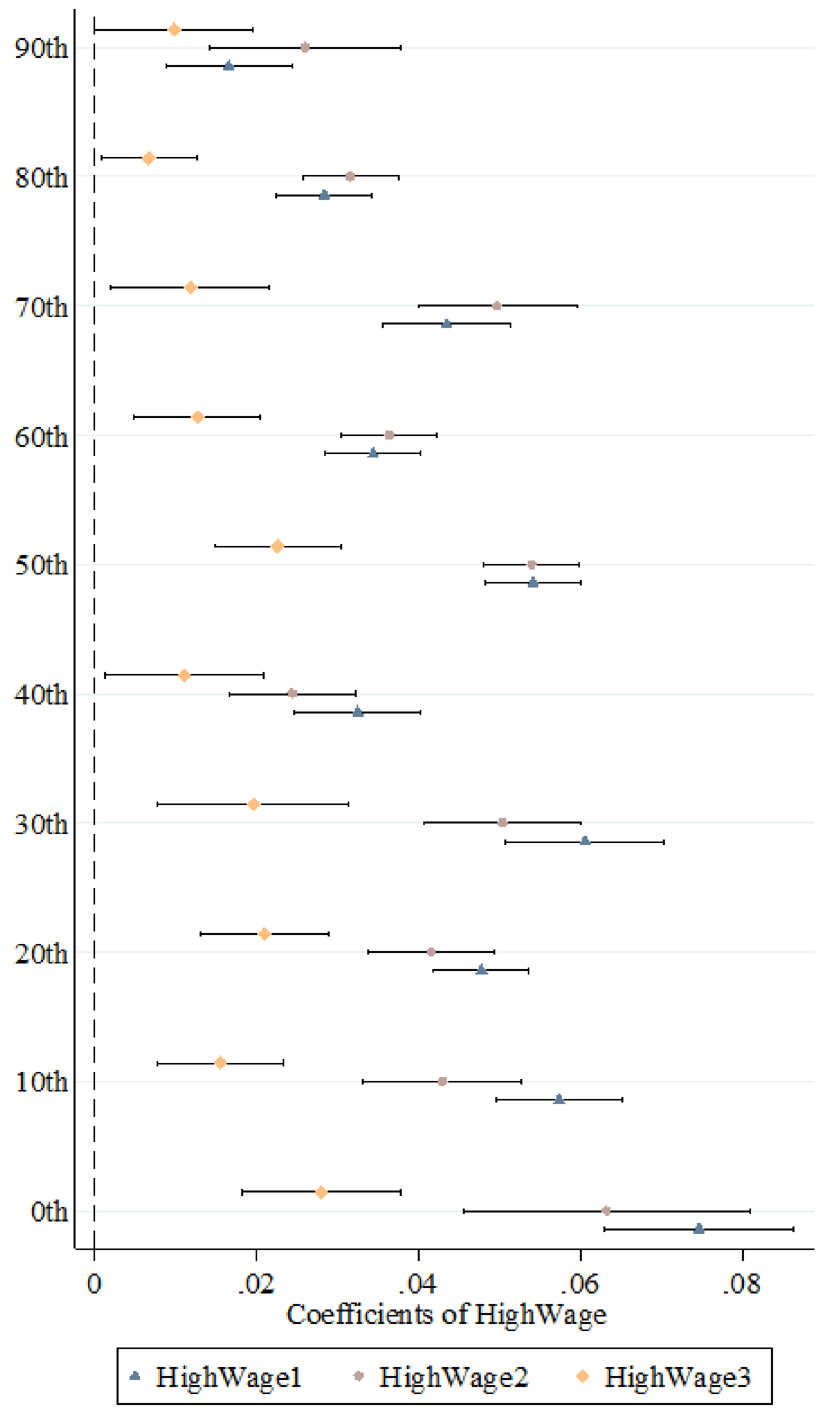


Figure B.11: Effects of HighWage by Education

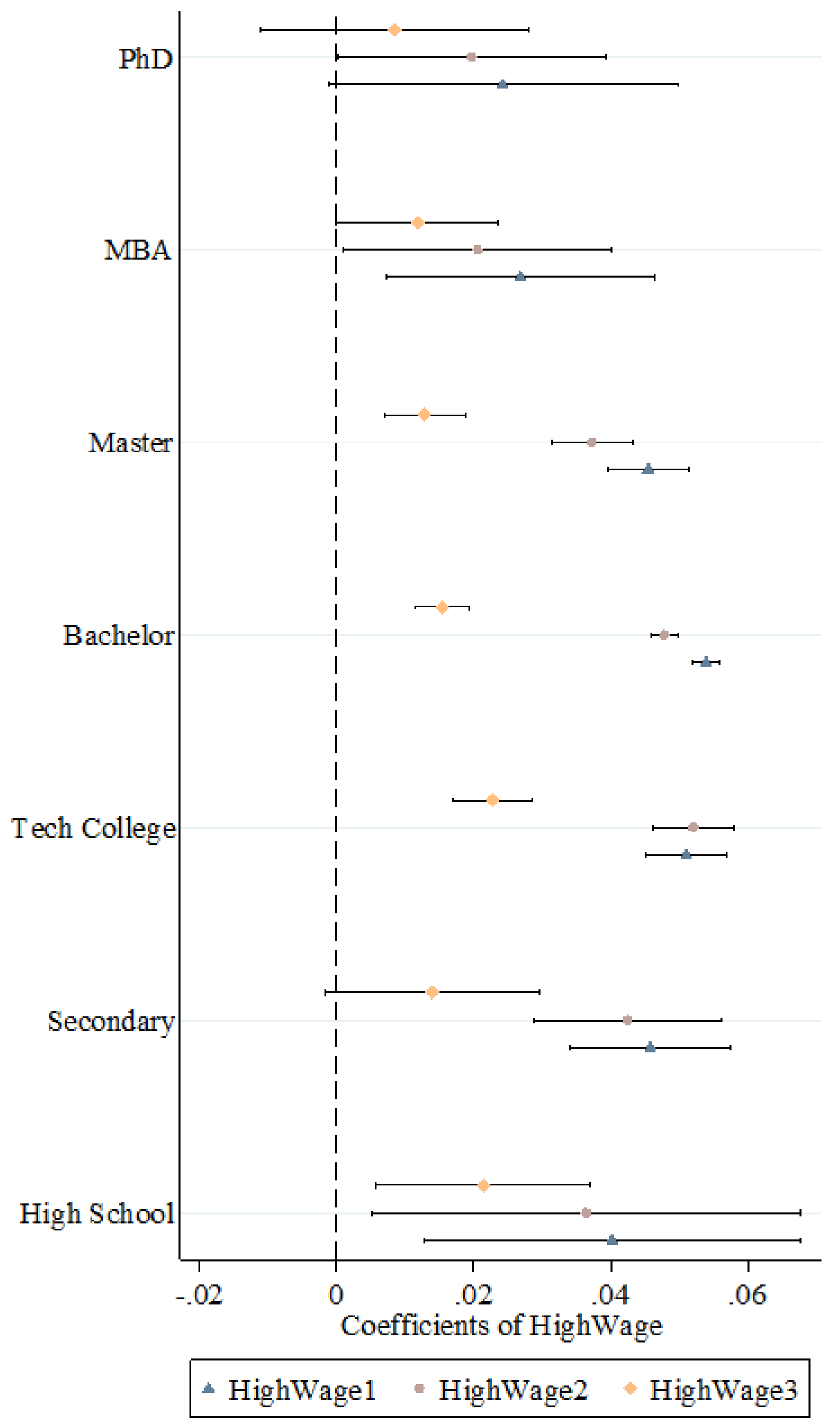


Table B.13: The Effect of High Wage Percentage on Wage Disclosure

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| **Panel A** | | | | | | | |
| HighWage1 | 0.0100*** | 0.0107*** | 0.0097*** | 0.0072*** | 0.0042*** | 0.0042*** | 0.0040*** |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Wage | | -0.0044*** | -0.0029*** | -0.0013*** | -0.0001*** | -0.0002*** | -0.0002*** |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Male | | 0.0820*** | 0.0936*** | 0.0724*** | 0.0822*** | 0.0865*** | 0.0873*** |
| | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| **Panel B** | | | | | | | |
| HighWage2 | 0.0098*** | 0.0106*** | 0.0094*** | 0.0073*** | 0.0059*** | 0.0049*** | 0.0044*** |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Wage | | -0.0044*** | -0.0032*** | -0.0013*** | -0.0007*** | -0.0004*** | -0.0005*** |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Male | | 0.0820*** | 0.0929*** | 0.0724*** | 0.0810*** | 0.0885*** | 0.0817*** |
| | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| **Panel C** | | | | | | | |
| HighWage3 | 0.0092*** | 0.0115*** | 0.0093*** | 0.0082*** | 0.0063*** | 0.0043*** | 0.0038*** |
| | (0.001) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Wage | | -0.0038*** | -0.0014*** | -0.0012*** | -0.0011*** | -0.0003*** | -0.0001*** |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Male | | 0.0813*** | 0.0893*** | 0.0739*** | 0.0870*** | 0.0864*** | 0.0852*** |
| | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Demographics | | | Yes | Yes | Yes | Yes | Yes |
| Edu & Exp | | | | Yes | Yes | Yes | Yes |
| Classification & Match | | | | | Yes | Yes | Yes |
| Location FE | | | | | | Yes | Yes |
| Industry FE | | | | | | | Yes |
| Occupation FE | | | | | | | Yes |
| 'Effective' *N* | 941,733 | 941,733 | 941,733 | 941,733 | 941,733 | 941,697 | 941,695 |

Standard errors in parentheses, clustered by worker's sub-occupation. *** p<0.01, ** p<0.05, * p<0.1

Note: The high-wage workers are measured by binary variables in Section 5. We further check the effect of the deviation from the worker's expected wage on wage revealing by using the percentage difference between worker's actual wage and his expected wage and replicating regressions in Table 1 with the continuous measure of HighWage as $HighWage_i = \frac{w_i - W_{Ei}}{W_{Ei}}$.

## B.5 Employer's Response on Success Rate

### B.5.1 Predicted Wage on Job Level

In Section 2.4.2, when employers infer the applicants' wages from their information on resume, the expected wage is the predicted wage based on the following regression in the sample of job applicants.

$$wage_j = \beta_0 + \beta_1 X_j + e_j \tag{B.15}$$

The covariates controlled in the regressions are: gender, age, marital status, years of working experience, years of education (converted from the highest degree), whether the worker is unemployed, and whether the worker is currently working in the job's industry and occupation.

Compared to predicted wage in worker level in Appendix B.4.1, the prediction on job level is more compact with fewer controls. There are two reasons for using the shorter regression: First, workers who apply to the same job share sufficient similarities such as the same industry and occupation, so the reduction of controls is not going to dramatically affect the accuracy of prediction. Second, some of jobs received only a few applicants, and a large set of controls are not allowed in the regression due to small sample size. In order to keep more observations in the analysis sample, we prefer the fewer controls in the prediction regression. In total, the regression is conducted on 83.7% of jobs, covering 92.4% of job applications.

In Table B.14, we list the average effects and standard errors of coefficients on the above covariates in the wage regressions on job level (the mean and standard error of the coefficients from 275,306 regressions).

Table B.14: Average Effects on Wage (Job Level)

| Variables | Average Coefficients |
|---|---|
| Male | 0.9084*** |
| | (0.235) |
| Age | 0.1347*** |
| | (0.040) |
| Single | 0.0643*** |
| | (0.019) |
| Married | 0.4095*** |
| | (0.126) |
| Experience | 1.2566*** |
| | (0.094) |
| Education | 0.3692*** |
| | (0.057) |
| Unemployed | -0.0677*** |
| | (0.021) |
| Same Industry | 0.6528*** |
| | (0.009) |
| Same Occupation | 0.8642*** |
| | (0.007) |
| Number of Jobs | 275,306 |

Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

### B.5.2 Extensions of Table 2.2

Table B.15: The Effect of Wage Disclosure on Save

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| | | | | **Panel A** | | | |
| OverWage1 | 0.0174*** | 0.0150*** | 0.0138*** | 0.0138*** | 0.0117*** | 0.0106*** | 0.0053*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.000) | (0.000) | (0.001) |
| Disclose* UnderWage1 | -0.0086*** | -0.0103*** | -0.0075*** | -0.0056*** | -0.0063*** | -0.0059*** | -0.0026* |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.001) |
| Disclose* OverWage1 | -0.0063*** | -0.0076*** | -0.0054*** | -0.0041*** | -0.0052*** | -0.0051*** | -0.0034** |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.001) |
| | | | | **Panel B** | | | |
| OverWage2 | 0.0122*** | 0.0114*** | 0.0094*** | 0.0108*** | 0.0095*** | 0.0093*** | 0.0040*** |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.001) |
| Disclose* UnderWage2 | -0.0070*** | -0.0090*** | -0.0061*** | -0.0043*** | -0.0053*** | -0.0053*** | -0.0021 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.001) |
| Disclose* OverWage2 | -0.0072*** | -0.0088*** | -0.0067*** | -0.0050*** | -0.0059*** | -0.0056*** | -0.0037*** |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.001) |
| | | | | **Panel C** | | | |
| OverWage3 | 0.0158*** | 0.0156*** | 0.0144*** | 0.0138*** | 0.0127*** | 0.0124*** | 0.0024*** |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.001) |
| Disclose* UnderWage3 | -0.0076*** | -0.0090*** | -0.0062*** | -0.0042*** | -0.0050*** | -0.0047*** | -0.0024* |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.001) |
| Disclose* OverWage3 | -0.0079*** | -0.0092*** | -0.0067*** | -0.0050*** | -0.0061*** | -0.0061*** | -0.0034** |
| | (0.001) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.001) |
| Job Characteristics | | Yes | Yes | Yes | Yes | | |
| Job Match | | | Yes | Yes | Yes | Yes | Yes |
| Worker Characteristics | | | | Yes | Yes | Yes | |
| City, Time, Ind & Occ FE | | | | | Yes | | |
| Firm FE | | | | | Yes | | |
| Job FE | | | | | | Yes | Yes |
| Worker FE | | | | | | | Yes |
| 'Effective' *N* | 3,542,049 | 3,542,049 | 3,542,049 | 3,542,049 | 3,541,342 | 3,486,460 | 3,201,726 |

Standard errors in parentheses, clustered by job. *** p<0.01, ** p<0.05, * p<0.1

Note: This table has the same identifications in Table 2. The outcome variable is resume being saved by hiring agents.

Table B.16: The Effect of Wage Disclosure on Suitable

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| | | | | **Panel A** | | | |
| OverWage1 | 0.0207*** | 0.0131*** | 0.0137*** | 0.0159*** | 0.0132*** | 0.0086*** | 0.0051*** |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.001) | (0.000) | (0.001) |
| Disclose* UnderWage1 | -0.0154*** | -0.0188*** | -0.0152*** | -0.0147*** | -0.0080*** | -0.0063*** | -0.0008 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.000) | (0.001) |
| Disclose* OverWage1 | -0.0081*** | -0.0123*** | -0.0096*** | -0.0094*** | -0.0057*** | -0.0046*** | -0.0013 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.000) | (0.001) |
| | | | | **Panel B** | | | |
| OverWage2 | 0.0076*** | 0.0083*** | 0.0072*** | 0.0109*** | 0.0087*** | 0.0074*** | 0.0037*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.000) | (0.001) |
| Disclose* UnderWage2 | -0.0150*** | -0.0185*** | -0.0147*** | -0.0145*** | -0.0081*** | -0.0062*** | -0.0006 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.000) | (0.002) |
| Disclose* OverWage2 | -0.0083*** | -0.0129*** | -0.0103*** | -0.0096*** | -0.0055*** | -0.0047*** | -0.0013 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.000) | (0.001) |
| | | | | **Panel C** | | | |
| OverWage3 | 0.0172*** | 0.0174*** | 0.0170*** | 0.0160*** | 0.0142*** | 0.0118*** | 0.0023*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.000) | (0.001) |
| Disclose* UnderWage3 | -0.0141*** | -0.0172*** | -0.0137*** | -0.0129*** | -0.0066*** | -0.0050*** | -0.0005 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.000) | (0.002) |
| Disclose* OverWage3 | -0.0092*** | -0.0137*** | -0.0109*** | -0.0103*** | -0.0063*** | -0.0056*** | -0.0013 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.000) | (0.002) |
| Job Characteristics | | Yes | Yes | Yes | Yes | | |
| Job Match | | | Yes | Yes | Yes | Yes | Yes |
| Worker Characteristics | | | | Yes | Yes | Yes | |
| City, Time, Ind & Occ FE | | | | | Yes | | |
| Firm FE | | | | | Yes | | |
| Job FE | | | | | | Yes | Yes |
| Worker FE | | | | | | | Yes |
| 'Effective' *N* | 3,542,049 | 3,542,049 | 3,542,049 | 3,542,049 | 3,541,342 | 3,486,460 | 3,201,726 |

Standard errors in parentheses, clustered by job. *** p<0.01, ** p<0.05, * p<0.1

Note: This table has the same identifications in Table 2. The outcome variable is application marked as suitable by hiring agents.

### B.5.3 Complete Version of Table 2.2

Table B.17: The Effect of Wage Disclosure on Becoming a Recruiting Target (Panel A)

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| OverWage1 | 0.0098*** | 0.0078*** | 0.0067*** | 0.0078*** | 0.0057*** | 0.0059*** | 0.0034*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.000) | (0.000) | (0.000) |
| Disclose* UnderWage1 | -0.0033*** | -0.0046*** | -0.0029*** | -0.0022*** | -0.0029*** | -0.0029*** | -0.0011 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.001) |
| Disclose* OverWage1 | -0.0029*** | -0.0043*** | -0.0030*** | -0.0025*** | -0.0029*** | -0.0031*** | -0.0023** |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.001) |
| Male | | -0.0009 | 0.0001 | 0.0009 | -0.0011*** | -0.0012*** | |
| | | (0.001) | (0.001) | (0.001) | (0.000) | (0.000) | |
| Job Wage lower bound | | -0.0003 | -0.0003 | -0.0003** | -0.0002*** | | |
| | | (0.000) | (0.000) | (0.000) | (0.000) | | |
| Job Wage upper bound | | 0.0001 | 0.0001 | 0.0002* | 0.0001** | | |
| | | (0.000) | (0.000) | (0.000) | (0.000) | | |
| Job Wage hidden | | -0.0017 | -0.0014 | -0.0010 | 0.0008 | | |
| | | (0.002) | (0.002) | (0.002) | (0.001) | | |
| Job Age lower bound | | 0.0009** | 0.0012*** | 0.0011*** | 0.0013*** | | |
| | | (0.000) | (0.000) | (0.000) | (0.000) | | |
| Job Age upper bound | | -0.0005** | -0.0006** | -0.0005* | -0.0007*** | | |
| | | (0.000) | (0.000) | (0.000) | (0.000) | | |
| Job experience | | -0.0008** | -0.0007** | -0.0005* | 0.0000 | | |
| | | (0.000) | (0.000) | (0.000) | (0.000) | | |
| Job Tongzhao Degree | | 0.0063*** | 0.0076*** | 0.0072*** | 0.0037*** | | |
| | | (0.002) | (0.002) | (0.002) | (0.001) | | |
| Job Subordinate # | | 0.0030** | 0.0029** | 0.0029** | 0.0002 | | |
| | | (0.001) | (0.001) | (0.001) | (0.001) | | |
| Job Feedback days | | 0.0015** | 0.0014** | 0.0013** | -0.0006*** | | |
| | | (0.001) | (0.001) | (0.001) | (0.000) | | |
| Job Post doc | | 0.2861*** | 0.3086*** | 0.3040*** | -0.0537 | | |
| | | (0.004) | (0.004) | (0.005) | (0.040) | | |
| Job Phd | | 0.0137 | 0.0288*** | 0.0255*** | 0.0266*** | | |
| | | (0.009) | (0.009) | (0.009) | (0.007) | | |
| Job MBA | | 0.0708 | 0.0880 | 0.0843 | 0.0837 | | |
| | | (0.064) | (0.062) | (0.062) | (0.065) | | |
| Job Master | | 0.0171** | 0.0232*** | 0.0213** | 0.0158*** | | |
| | | (0.008) | (0.008) | (0.009) | (0.003) | | |
| Job Bachelor | | -0.0057 | -0.0043 | -0.0046 | 0.0053*** | | |
| | | (0.004) | (0.004) | (0.004) | (0.002) | | |
| Job Tech college | | 0.0095** | 0.0057 | 0.0051 | 0.0023 | | |
| | | (0.004) | (0.004) | (0.004) | (0.002) | | |
| Job Secondary | | 0.0508*** | 0.0466*** | 0.0440*** | -0.0046 | | |
| | | (0.011) | (0.011) | (0.012) | (0.005) | | |
| Job High School | | 0.5661*** | 0.5598*** | 0.5569*** | -0.0849 | | |
| | | (0.042) | (0.042) | (0.042) | (0.059) | | |
| Worker over edu | | | 0.0048*** | 0.0048*** | 0.0037*** | 0.0037*** | 0.0044*** |

| | | | | | |
|---|---|---|---|---|---|
| | (0.001) | (0.001) | (0.000) | (0.000) | (0.001) |
| Worker under edu | -0.0218*** | -0.0193*** | -0.0192*** | -0.0201*** | -0.0162*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Mis tongzhao degree | -0.0119*** | -0.0106*** | -0.0092*** | -0.0091*** | -0.0032*** |
| | (0.001) | (0.001) | (0.000) | (0.000) | (0.001) |
| Worker over age | -0.0279*** | -0.0242*** | -0.0262*** | -0.0286*** | -0.0135*** |
| | (0.003) | (0.003) | (0.002) | (0.002) | (0.002) |
| Worker under age | -0.0127*** | -0.0130*** | -0.0167*** | -0.0125*** | -0.0107*** |
| | (0.004) | (0.005) | (0.003) | (0.003) | (0.003) |
| Mis gender | -0.0292*** | -0.0291*** | -0.0329*** | -0.0347*** | -0.0289*** |
| | (0.004) | (0.004) | (0.004) | (0.005) | (0.004) |
| Worker under exp | -0.0074*** | -0.0085*** | -0.0088*** | -0.0093*** | -0.0115*** |
| | (0.001) | (0.001) | (0.000) | (0.000) | (0.001) |
| City: job = desired | -0.0047*** | -0.0047*** | -0.0007 | -0.0005 | -0.0003 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| City: job = now | -0.0026*** | -0.0017** | 0.0010* | 0.0007 | 0.0017* |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Province: job = desired | 0.0039*** | 0.0027*** | 0.0019*** | 0.0020*** | 0.0019** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Province: job = now | 0.0040*** | 0.0028*** | 0.0038*** | 0.0046*** | 0.0015* |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Main ind: job = desired | -0.0051*** | -0.0018 | 0.0050*** | 0.0067*** | 0.0050*** |
| | (0.001) | (0.002) | (0.000) | (0.000) | (0.000) |
| Main ind: job = now | 0.0112*** | 0.0087*** | 0.0054*** | 0.0060*** | 0.0039*** |
| | (0.002) | (0.001) | (0.000) | (0.000) | (0.001) |
| Sub ind: job = desired | -0.0019* | -0.0039*** | -0.0004 | -0.0005 | -0.0009** |
| | (0.001) | (0.001) | (0.000) | (0.000) | (0.000) |
| Sub ind: job = now | 0.0030** | 0.0043*** | 0.0070*** | 0.0074*** | 0.0053*** |
| | (0.001) | (0.001) | (0.000) | (0.000) | (0.001) |
| Main occ: job = desired | 0.0062*** | 0.0049*** | 0.0022*** | 0.0026*** | 0.0024*** |
| | (0.001) | (0.001) | (0.000) | (0.000) | (0.000) |
| Main occ: job = now | 0.0075*** | 0.0076*** | 0.0076*** | 0.0080*** | 0.0051*** |
| | (0.001) | (0.001) | (0.000) | (0.000) | (0.000) |
| Sub occ: job = desired | 0.0023*** | 0.0016** | 0.0031*** | 0.0033*** | 0.0020*** |
| | (0.001) | (0.001) | (0.000) | (0.000) | (0.000) |
| Sub occ: job = now | 0.0026*** | 0.0021*** | 0.0041*** | 0.0046*** | 0.0041*** |
| | (0.001) | (0.001) | (0.000) | (0.000) | (0.000) |
| Single | | 0.0103*** | 0.0007 | 0.0001 | |
| | | (0.001) | (0.000) | (0.000) | |
| Married | | 0.0092*** | -0.0003 | -0.0008* | |
| | | (0.001) | (0.000) | (0.000) | |
| Intensively Search | | 0.0050*** | 0.0053*** | 0.0053*** | |
| | | (0.000) | (0.000) | (0.000) | |
| Moderately Search | | 0.0069*** | 0.0064*** | 0.0063*** | |
| | | (0.000) | (0.000) | (0.000) | |
| Stay | | 0.0200*** | 0.0196*** | 0.0152** | |
| | | (0.008) | (0.007) | (0.007) | |
| 985/211 school | | 0.0006 | 0.0010*** | 0.0010*** | |

| | | | | |
|---|---|---|---|---|
| | (0.001) | (0.000) | (0.000) | |
| Tenure 1 | -0.0007*** | -0.0010*** | -0.0009*** | |
| | (0.000) | (0.000) | (0.000) | |
| Tenure 2 | -0.0003*** | -0.0006*** | -0.0006*** | |
| | (0.000) | (0.000) | (0.000) | |
| Domestic 1-19 | 0.0016** | 0.0021*** | 0.0022*** | |
| | (0.001) | (0.001) | (0.000) | |
| Domestic 20-39 | 0.0013* | 0.0016*** | 0.0015*** | |
| | (0.001) | (0.001) | (0.001) | |
| Domestic 40-59 | 0.0022** | 0.0018*** | 0.0017*** | |
| | (0.001) | (0.001) | (0.001) | |
| Domestic 60-99 | 0.0024*** | 0.0025*** | 0.0027*** | |
| | (0.001) | (0.000) | (0.000) | |
| Domestic 100-199 | 0.0016*** | 0.0022*** | 0.0023*** | |
| | (0.000) | (0.000) | (0.000) | |
| World 1-19 | 0.0161*** | 0.0183*** | 0.0186*** | |
| | (0.005) | (0.004) | (0.004) | |
| World 20-39 | 0.0046 | 0.0057*** | 0.0059*** | |
| | (0.003) | (0.002) | (0.002) | |
| World 40-59 | -0.0031** | -0.0006 | -0.0007 | |
| | (0.001) | (0.001) | (0.001) | |
| World 60-99 | 0.0014 | 0.0002 | 0.0003 | |
| | (0.002) | (0.001) | (0.001) | |
| World 100-199 | -0.0004 | -0.0004 | 0.0001 | |
| | (0.001) | (0.001) | (0.001) | |
| Days of Account | -0.0001*** | -0.0001*** | -0.0001*** | |
| | (0.000) | (0.000) | (0.000) | |
| Resume Completeness | -0.0022*** | -0.0008*** | -0.0006*** | |
| | (0.000) | (0.000) | (0.000) | |
| Membership | -0.0017*** | -0.0012*** | -0.0010*** | |
| | (0.000) | (0.000) | (0.000) | |
| Elite | -0.0044*** | -0.0007 | -0.0005 | |
| | (0.001) | (0.001) | (0.001) | |
| Batch Apply | -0.0169*** | -0.0140*** | -0.0122*** | -0.0050*** |
| | (0.001) | (0.000) | (0.000) | (0.000) |
| Lens*1 | 0.0227*** | 0.0198*** | 0.0185*** | 0.0159*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| Lens*2 | 0.0241*** | 0.0216*** | 0.0199*** | 0.0211*** |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| Number of Application | -0.0000 | -0.0000*** | | |
| | (0.000) | (0.000) | | |
| Time, Location FE | | Yes | | |
| Job Ind & Occ FE | | Yes | | |
| Firm FE | | Yes | | |
| Job FE | | | Yes | Yes |
| Worker FE | | | | Yes |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 'Effective' $N$ | 3,542,049 | 3,542,049 | 3,542,049 | 3,542,049 | 3,541,342 | 3,486,460 | 3,201,726 |
| $R^2$ | 0.001 | 0.003 | 0.007 | 0.009 | 0.205 | 0.350 | 0.484 |

Standard errors in parentheses, clustered by job. *** p<0.01, ** p<0.05, * p<0.1

Table B.18: The Effect of Wage Disclosure on Becoming a Recruiting Target (Panel B)

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| OverWage2 | 0.0057*** | 0.0054*** | 0.0042*** | 0.0054*** | 0.0046*** | 0.0050*** | 0.0024*** |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Disclose* UnderWage2 | -0.0032*** | -0.0047*** | -0.0029*** | -0.0024*** | -0.0029*** | -0.0028*** | -0.0011 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.001) |
| Disclose* OverWage2 | -0.0028*** | -0.0043*** | -0.0030*** | -0.0023*** | -0.0029*** | -0.0031*** | -0.0022** |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.001) |
| Male | | -0.0009 | 0.0002 | 0.0009 | -0.0011*** | -0.0012*** | |
| | | (0.001) | (0.001) | (0.001) | (0.000) | (0.000) | |
| Job Wage lower bound | | -0.0003** | -0.0003** | -0.0004*** | -0.0003*** | | |
| | | (0.000) | (0.000) | (0.000) | (0.000) | | |
| Job Wage upper bound | | 0.0001 | 0.0001 | 0.0002* | 0.0001** | | |
| | | (0.000) | (0.000) | (0.000) | (0.000) | | |
| Job Wage hidden | | -0.0014 | -0.0011 | -0.0007 | 0.0009 | | |
| | | (0.002) | (0.002) | (0.002) | (0.001) | | |
| Job Age lower bound | | 0.0010** | 0.0013*** | 0.0011*** | 0.0013*** | | |
| | | (0.000) | (0.000) | (0.000) | (0.000) | | |
| Job Age upper bound | | -0.0005** | -0.0006** | -0.0005* | -0.0007*** | | |
| | | (0.000) | (0.000) | (0.000) | (0.000) | | |
| Job experience | | -0.0007* | -0.0007* | -0.0005 | 0.0000 | | |
| | | (0.000) | (0.000) | (0.000) | (0.000) | | |
| Job Tongzhao Degree | | 0.0063*** | 0.0075*** | 0.0072*** | 0.0037*** | | |
| | | (0.002) | (0.002) | (0.002) | (0.001) | | |
| Job Subordinate # | | 0.0031** | 0.0030** | 0.0030** | 0.0003 | | |
| | | (0.001) | (0.001) | (0.001) | (0.001) | | |
| Job Feedback days | | 0.0015** | 0.0014** | 0.0013** | -0.0006*** | | |
| | | (0.001) | (0.001) | (0.001) | (0.000) | | |
| Job Post doc | | 0.2837*** | 0.3066*** | 0.3015*** | -0.0563 | | |
| | | (0.004) | (0.004) | (0.005) | (0.041) | | |
| Job Phd | | 0.0134 | 0.0287*** | 0.0252*** | 0.0264*** | | |
| | | (0.009) | (0.009) | (0.009) | (0.007) | | |
| Job MBA | | 0.0717 | 0.0887 | 0.0850 | 0.0840 | | |
| | | (0.065) | (0.063) | (0.062) | (0.065) | | |
| Job Master | | 0.0168** | 0.0231*** | 0.0211** | 0.0158*** | | |
| | | (0.008) | (0.008) | (0.009) | (0.003) | | |
| Job Bachelor | | -0.0059 | -0.0045 | -0.0048 | 0.0053*** | | |
| | | (0.004) | (0.004) | (0.004) | (0.002) | | |
| Job Tech college | | 0.0096** | 0.0057 | 0.0051 | 0.0023 | | |
| | | (0.004) | (0.004) | (0.004) | (0.002) | | |
| Job Secondary | | 0.0512*** | 0.0468*** | 0.0442*** | -0.0045 | | |
| | | (0.011) | (0.011) | (0.012) | (0.005) | | |
| Job High School | | 0.5679*** | 0.5612*** | 0.5586*** | -0.0833 | | |
| | | (0.042) | (0.042) | (0.042) | (0.059) | | |
| Worker over edu | | | 0.0050*** | 0.0050*** | 0.0037*** | 0.0037*** | 0.0044*** |

| | | | | | |
|---|---|---|---|---|---|
| | (0.001) | (0.001) | (0.000) | (0.000) | (0.001) |
| Worker under edu | -0.0218*** | -0.0193*** | -0.0192*** | -0.0201*** | -0.0162*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Mis tongzhao degree | -0.0118*** | -0.0106*** | -0.0092*** | -0.0091*** | -0.0032*** |
| | (0.001) | (0.001) | (0.000) | (0.000) | (0.001) |
| Worker over age | -0.0275*** | -0.0239*** | -0.0262*** | -0.0286*** | -0.0135*** |
| | (0.003) | (0.003) | (0.002) | (0.002) | (0.002) |
| Worker under age | -0.0130*** | -0.0132*** | -0.0167*** | -0.0124*** | -0.0107*** |
| | (0.004) | (0.005) | (0.003) | (0.003) | (0.003) |
| Mis gender | -0.0292*** | -0.0291*** | -0.0329*** | -0.0346*** | -0.0289*** |
| | (0.004) | (0.004) | (0.004) | (0.005) | (0.004) |
| Worker under exp | -0.0079*** | -0.0089*** | -0.0090*** | -0.0093*** | -0.0114*** |
| | (0.001) | (0.001) | (0.000) | (0.000) | (0.001) |
| City: job = desired | -0.0047*** | -0.0047*** | -0.0007 | -0.0004 | -0.0003 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| City: job = now | -0.0024*** | -0.0015* | 0.0010* | 0.0007 | 0.0017* |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Province: job = desired | 0.0039*** | 0.0026*** | 0.0019*** | 0.0020*** | 0.0019** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Province: job = now | 0.0039*** | 0.0028*** | 0.0038*** | 0.0046*** | 0.0015* |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Main ind: job = desired | -0.0053*** | -0.0020 | 0.0050*** | 0.0067*** | 0.0050*** |
| | (0.001) | (0.002) | (0.000) | (0.000) | (0.000) |
| Main ind: job = now | 0.0113*** | 0.0087*** | 0.0054*** | 0.0060*** | 0.0039*** |
| | (0.002) | (0.001) | (0.000) | (0.000) | (0.001) |
| Sub ind: job = desired | -0.0017* | -0.0038*** | -0.0004 | -0.0005 | -0.0009** |
| | (0.001) | (0.001) | (0.000) | (0.000) | (0.000) |
| Sub ind: job = now | 0.0031** | 0.0043*** | 0.0071*** | 0.0074*** | 0.0053*** |
| | (0.001) | (0.001) | (0.000) | (0.000) | (0.001) |
| Main occ: job = desired | 0.0062*** | 0.0050*** | 0.0022*** | 0.0026*** | 0.0024*** |
| | (0.001) | (0.001) | (0.000) | (0.000) | (0.000) |
| Main occ: job = now | 0.0074*** | 0.0075*** | 0.0076*** | 0.0081*** | 0.0051*** |
| | (0.001) | (0.001) | (0.000) | (0.000) | (0.000) |
| Sub occ: job = desired | 0.0023*** | 0.0017** | 0.0031*** | 0.0033*** | 0.0020*** |
| | (0.001) | (0.001) | (0.000) | (0.000) | (0.000) |
| Sub occ: job = now | 0.0026*** | 0.0021*** | 0.0041*** | 0.0046*** | 0.0041*** |
| | (0.001) | (0.001) | (0.000) | (0.000) | (0.000) |
| Single | | 0.0105*** | 0.0007 | 0.0001 | |
| | | (0.001) | (0.000) | (0.000) | |
| Married | | 0.0095*** | -0.0002 | -0.0008* | |
| | | (0.001) | (0.000) | (0.000) | |
| Intensively Search | | 0.0049*** | 0.0053*** | 0.0053*** | |
| | | (0.000) | (0.000) | (0.000) | |
| Moderately Search | | 0.0069*** | 0.0064*** | 0.0063*** | |
| | | (0.000) | (0.000) | (0.000) | |
| Stay | | 0.0201*** | 0.0197*** | 0.0153** | |
| | | (0.008) | (0.007) | (0.007) | |
| 985/211 school | | 0.0007 | 0.0010*** | 0.0010*** | |

| | | | | |
|---|---|---|---|---|
| | (0.001) | (0.000) | (0.000) | |
| Tenure 1 | -0.0007*** | -0.0010*** | -0.0010*** | |
| | (0.000) | (0.000) | (0.000) | |
| Tenure 2 | -0.0002*** | -0.0006*** | -0.0006*** | |
| | (0.000) | (0.000) | (0.000) | |
| Domestic 1-19 | 0.0016** | 0.0021*** | 0.0022*** | |
| | (0.001) | (0.001) | (0.000) | |
| Domestic 20-39 | 0.0012* | 0.0015*** | 0.0015*** | |
| | (0.001) | (0.001) | (0.001) | |
| Domestic 40-59 | 0.0021** | 0.0018*** | 0.0017*** | |
| | (0.001) | (0.001) | (0.001) | |
| Domestic 60-99 | 0.0024*** | 0.0025*** | 0.0027*** | |
| | (0.001) | (0.000) | (0.000) | |
| Domestic 100-199 | 0.0015*** | 0.0022*** | 0.0023*** | |
| | (0.000) | (0.000) | (0.000) | |
| World 1-19 | 0.0162*** | 0.0183*** | 0.0186*** | |
| | (0.005) | (0.004) | (0.004) | |
| World 20-39 | 0.0047* | 0.0057*** | 0.0060*** | |
| | (0.003) | (0.002) | (0.002) | |
| World 40-59 | -0.0028* | -0.0006 | -0.0007 | |
| | (0.001) | (0.001) | (0.001) | |
| World 60-99 | 0.0017 | 0.0002 | 0.0004 | |
| | (0.002) | (0.001) | (0.001) | |
| World 100-199 | -0.0003 | -0.0004 | 0.0001 | |
| | (0.001) | (0.001) | (0.001) | |
| Days of Account | -0.0000* | -0.0001*** | -0.0001*** | |
| | (0.000) | (0.000) | (0.000) | |
| Resume Completeness | -0.0022*** | -0.0008*** | -0.0006*** | |
| | (0.000) | (0.000) | (0.000) | |
| Membership | -0.0015*** | -0.0011*** | -0.0010*** | |
| | (0.000) | (0.000) | (0.000) | |
| Elite | -0.0044*** | -0.0007 | -0.0005 | |
| | (0.001) | (0.001) | (0.001) | |
| Batch Apply | -0.0169*** | -0.0140*** | -0.0122*** | -0.0050*** |
| | (0.001) | (0.000) | (0.000) | (0.000) |
| Lens*1 | 0.0226*** | 0.0198*** | 0.0185*** | 0.0159*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| Lens*2 | 0.0240*** | 0.0216*** | 0.0199*** | 0.0210*** |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| Number of Application | -0.0000 | -0.0000*** | | |
| | (0.000) | (0.000) | | |
| Time, Location FE | | Yes | | |
| Job Ind & Occ FE | | Yes | | |
| Firm FE | | Yes | | |
| Job FE | | | Yes | Yes |
| Worker FE | | | | Yes |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| ‘Effective’ *N* | 3,542,049 | 3,542,049 | 3,542,049 | 3,542,049 | 3,541,342 | 3,486,460 | 3,201,726 |
| $R^2$ | 0.000 | 0.003 | 0.006 | 0.009 | 0.205 | 0.350 | 0.484 |

Standard errors in parentheses, clustered by job. *** p<0.01, ** p<0.05, * p<0.1

Table B.19: The Effect of Wage Disclosure on Becoming a Recruiting Target (Panel C)

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| OverWage3 | 0.0073*** | 0.0072*** | 0.0067*** | 0.0069*** | 0.0066*** | 0.0069*** | 0.0019*** |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Disclose* UnderWage3 | -0.0037*** | -0.0048*** | -0.0031*** | -0.0023*** | -0.0027*** | -0.0025*** | -0.0014 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.001) |
| Disclose* OverWage3 | -0.0035*** | -0.0047*** | -0.0032*** | -0.0025*** | -0.0032*** | -0.0034*** | -0.0025** |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.001) |
| Male | | -0.0007 | 0.0005 | 0.0012* | -0.0004 | -0.0006* | |
| | | (0.001) | (0.001) | (0.001) | (0.000) | (0.000) | |
| Job Wage lower bound | | -0.0004** | -0.0004** | -0.0004*** | -0.0003*** | | |
| | | (0.000) | (0.000) | (0.000) | (0.000) | | |
| Job Wage upper bound | | 0.0002 | 0.0001 | 0.0002* | 0.0001** | | |
| | | (0.000) | (0.000) | (0.000) | (0.000) | | |
| Job Wage hidden | | -0.0004 | -0.0001 | 0.0001 | 0.0008 | | |
| | | (0.002) | (0.002) | (0.002) | (0.001) | | |
| Job Age lower bound | | 0.0009** | 0.0011** | 0.0010** | 0.0014*** | | |
| | | (0.000) | (0.000) | (0.000) | (0.000) | | |
| Job Age upper bound | | -0.0005* | -0.0005* | -0.0004 | -0.0008*** | | |
| | | (0.000) | (0.000) | (0.000) | (0.000) | | |
| Job experience | | -0.0006 | -0.0006 | -0.0004 | 0.0001 | | |
| | | (0.000) | (0.000) | (0.000) | (0.000) | | |
| Job Tongzhao Degree | | 0.0068*** | 0.0079*** | 0.0076*** | 0.0039*** | | |
| | | (0.002) | (0.002) | (0.002) | (0.001) | | |
| Job Subordinate # | | 0.0031** | 0.0030** | 0.0029** | 0.0003 | | |
| | | (0.001) | (0.001) | (0.001) | (0.001) | | |
| Job Feedback days | | 0.0016** | 0.0016** | 0.0015** | -0.0006*** | | |
| | | (0.001) | (0.001) | (0.001) | (0.000) | | |
| Job Post doc | | - | - | - | - | | |
| Job Phd | | 0.0098 | 0.0268*** | 0.0235** | 0.0251*** | | |
| | | (0.010) | (0.010) | (0.010) | (0.008) | | |
| Job MBA | | 0.0811 | 0.0978 | 0.0940 | 0.0861 | | |
| | | (0.068) | (0.065) | (0.065) | (0.068) | | |
| Job Master | | 0.0200** | 0.0260*** | 0.0236** | 0.0166*** | | |
| | | (0.009) | (0.009) | (0.009) | (0.003) | | |
| Job Bachelor | | -0.0032 | -0.0022 | -0.0027 | 0.0061*** | | |
| | | (0.005) | (0.005) | (0.005) | (0.002) | | |
| Job Tech college | | 0.0113** | 0.0065 | 0.0062 | 0.0023 | | |
| | | (0.005) | (0.005) | (0.005) | (0.002) | | |
| Job Secondary | | 0.0663*** | 0.0599*** | 0.0583*** | -0.0017 | | |
| | | (0.015) | (0.015) | (0.016) | (0.007) | | |
| Job High School | | 0.5687*** | 0.5613*** | 0.5590*** | -0.1142* | | |
| | | (0.044) | (0.044) | (0.044) | (0.060) | | |
| Worker over edu | | | 0.0059*** | 0.0056*** | 0.0043*** | 0.0044*** | 0.0041*** |

| | | | | | |
|---|---|---|---|---|---|
| | (0.001) | (0.001) | (0.000) | (0.000) | (0.001) |
| Worker under edu | -0.0219*** | -0.0192*** | -0.0191*** | -0.0199*** | -0.0152*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Mis tongzhao degree | -0.0111*** | -0.0101*** | -0.0087*** | -0.0088*** | -0.0032*** |
| | (0.001) | (0.001) | (0.001) | (0.000) | (0.001) |
| Worker over age | -0.0257*** | -0.0226*** | -0.0249*** | -0.0277*** | -0.0134*** |
| | (0.003) | (0.003) | (0.002) | (0.003) | (0.002) |
| Worker under age | -0.0140*** | -0.0140*** | -0.0177*** | -0.0137*** | -0.0116*** |
| | (0.005) | (0.005) | (0.003) | (0.003) | (0.003) |
| Mis gender | -0.0286*** | -0.0284*** | -0.0328*** | -0.0344*** | -0.0282*** |
| | (0.004) | (0.004) | (0.005) | (0.006) | (0.004) |
| Worker under exp | -0.0088*** | -0.0095*** | -0.0092*** | -0.0097*** | -0.0113*** |
| | (0.001) | (0.001) | (0.000) | (0.000) | (0.001) |
| City: job = desired | -0.0046*** | -0.0046*** | -0.0011** | -0.0008 | -0.0002 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| City: job = now | -0.0015 | -0.0008 | 0.0013** | 0.0007 | 0.0017* |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Province: job = desired | 0.0040*** | 0.0029*** | 0.0024*** | 0.0022*** | 0.0015 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Province: job = now | 0.0035*** | 0.0024** | 0.0037*** | 0.0046*** | 0.0017* |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Main ind: job = desired | -0.0049*** | -0.0018 | 0.0051*** | 0.0066*** | 0.0051*** |
| | (0.001) | (0.002) | (0.000) | (0.000) | (0.000) |
| Main ind: job = now | 0.0109*** | 0.0085*** | 0.0051*** | 0.0056*** | 0.0036*** |
| | (0.002) | (0.001) | (0.000) | (0.000) | (0.001) |
| Sub ind: job = desired | -0.0021* | -0.0040*** | -0.0006 | -0.0006* | -0.0009** |
| | (0.001) | (0.001) | (0.000) | (0.000) | (0.000) |
| Sub ind: job = now | 0.0029** | 0.0039*** | 0.0073*** | 0.0077*** | 0.0056*** |
| | (0.001) | (0.001) | (0.001) | (0.000) | (0.001) |
| Main occ: job = desired | 0.0064*** | 0.0052*** | 0.0019*** | 0.0023*** | 0.0021*** |
| | (0.001) | (0.001) | (0.000) | (0.000) | (0.000) |
| Main occ: job = now | 0.0075*** | 0.0076*** | 0.0074*** | 0.0079*** | 0.0049*** |
| | (0.001) | (0.001) | (0.000) | (0.000) | (0.000) |
| Sub occ: job = desired | 0.0023** | 0.0017** | 0.0031*** | 0.0031*** | 0.0019*** |
| | (0.001) | (0.001) | (0.000) | (0.000) | (0.000) |
| Sub occ: job = now | 0.0020*** | 0.0017** | 0.0037*** | 0.0044*** | 0.0040*** |
| | (0.001) | (0.001) | (0.000) | (0.000) | (0.000) |
| Single | | 0.0101*** | 0.0004 | -0.0001 | |
| | | (0.001) | (0.000) | (0.000) | |
| Married | | 0.0095*** | -0.0002 | -0.0007 | |
| | | (0.001) | (0.000) | (0.000) | |
| Intensively Search | | 0.0052*** | 0.0053*** | 0.0053*** | |
| | | (0.000) | (0.000) | (0.000) | |
| Moderately Search | | 0.0069*** | 0.0064*** | 0.0063*** | |
| | | (0.000) | (0.000) | (0.000) | |
| Stay | | 0.0224*** | 0.0206*** | 0.0153** | |
| | | (0.008) | (0.007) | (0.007) | |
| 985/211 school | | 0.0009* | 0.0013*** | 0.0012*** | |

| | | | | |
|---|---|---|---|---|
| | (0.001) | (0.000) | (0.000) | |
| Tenure 1 | -0.0005*** | -0.0009*** | -0.0008*** | |
| | (0.000) | (0.000) | (0.000) | |
| Tenure 2 | -0.0002*** | -0.0005*** | -0.0005*** | |
| | (0.000) | (0.000) | (0.000) | |
| Domestic 1-19 | 0.0019*** | 0.0021*** | 0.0023*** | |
| | (0.001) | (0.001) | (0.001) | |
| Domestic 20-39 | 0.0011 | 0.0012** | 0.0014*** | |
| | (0.001) | (0.001) | (0.001) | |
| Domestic 40-59 | 0.0020** | 0.0017** | 0.0016** | |
| | (0.001) | (0.001) | (0.001) | |
| Domestic 60-99 | 0.0026*** | 0.0025*** | 0.0026*** | |
| | (0.001) | (0.000) | (0.000) | |
| Domestic 100-199 | 0.0016*** | 0.0021*** | 0.0022*** | |
| | (0.000) | (0.000) | (0.000) | |
| World 1-19 | 0.0164*** | 0.0186*** | 0.0187*** | |
| | (0.005) | (0.004) | (0.004) | |
| World 20-39 | 0.0039 | 0.0046** | 0.0054*** | |
| | (0.003) | (0.002) | (0.002) | |
| World 40-59 | -0.0028* | -0.0003 | -0.0006 | |
| | (0.002) | (0.001) | (0.001) | |
| World 60-99 | 0.0019 | 0.0004 | 0.0003 | |
| | (0.002) | (0.001) | (0.001) | |
| World 100-199 | -0.0004 | -0.0003 | 0.0000 | |
| | (0.001) | (0.001) | (0.001) | |
| Days of Account | -0.0000 | -0.0001*** | -0.0001*** | |
| | (0.000) | (0.000) | (0.000) | |
| Resume Completeness | -0.0021*** | -0.0008*** | -0.0007*** | |
| | (0.000) | (0.000) | (0.000) | |
| Membership | -0.0015*** | -0.0011*** | -0.0010*** | |
| | (0.000) | (0.000) | (0.000) | |
| Elite | -0.0029** | 0.0000 | 0.0001 | |
| | (0.001) | (0.001) | (0.001) | |
| Batch Apply | -0.0157*** | -0.0126*** | -0.0114*** | -0.0045*** |
| | (0.001) | (0.000) | (0.000) | (0.000) |
| Lens*1 | 0.0216*** | 0.0188*** | 0.0179*** | 0.0158*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| Lens*2 | 0.0238*** | 0.0213*** | 0.0200*** | 0.0205*** |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| Number of Application | -0.0000 | -0.0000*** | | |
| | (0.000) | (0.000) | | |
| Time, Location FE | | Yes | | |
| Job Ind & Occ FE | | Yes | | |
| Firm FE | | Yes | | |
| Job FE | | | Yes | Yes |
| Worker FE | | | | Yes |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 'Effective' *N* | 3,263,843 | 3,263,843 | 3,263,843 | 3,263,843 | 3,263,649 | 3,254,545 | 2,987,419 |
| $R^2$ | 0.000 | 0.003 | 0.007 | 0.009 | 0.215 | 0.334 | 0.471 |

Standard errors in parentheses, clustered by job. *** p<0.01, ** p<0.05, * p<0.1

### B.5.4 Robustness Check on Employer's Response

Figure B.12: Effects of OverWage on Target by Firm Size

Figure B.12a: Coefficients of Overwage on Target

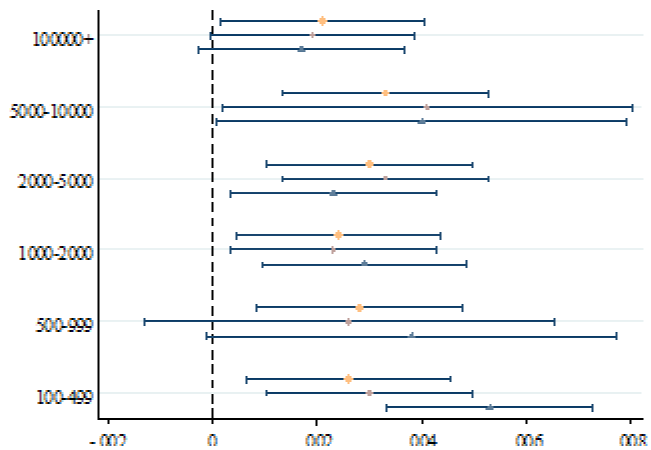



Figure B.12b: Coefficients of Underwage*Disclose on Target

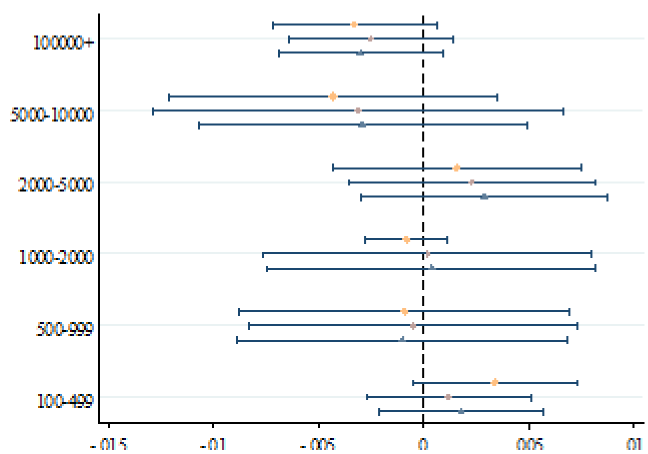


Figure B.12c: Coefficients of Overwage*Disclose on Target

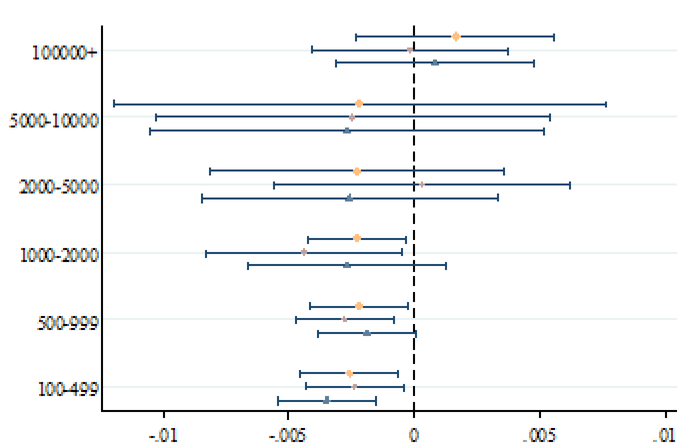



Figure B.13: Effects of OverWage on Target by Industry Tightness

Figure B.13a: Coefficients of Overwage on Target

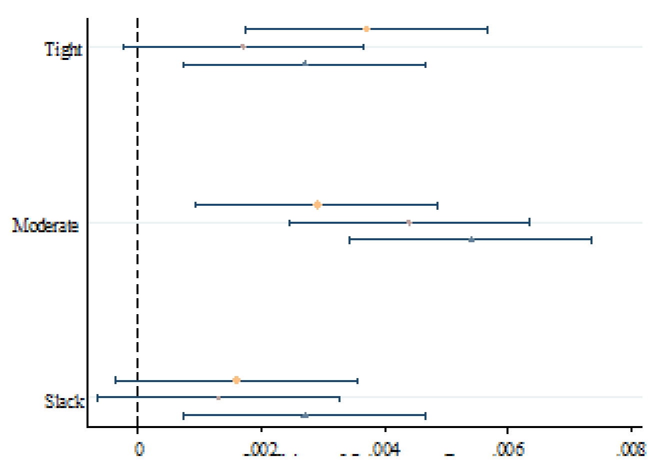


Figure B.13b: Coefficients of Underwage*Disclose on Target

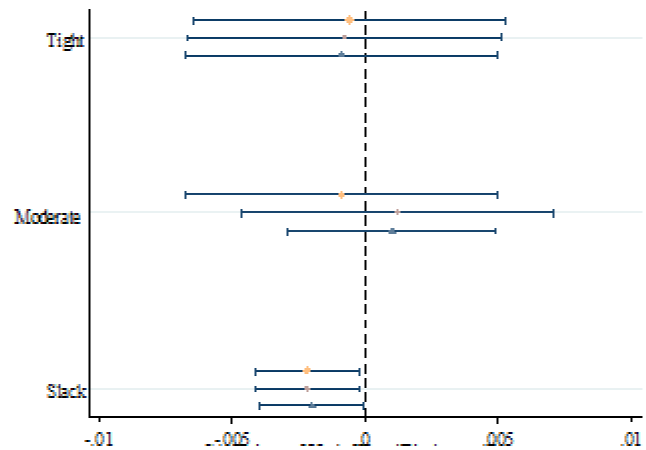



Figure B.13c: Coefficients of Overwage*Disclose on Target

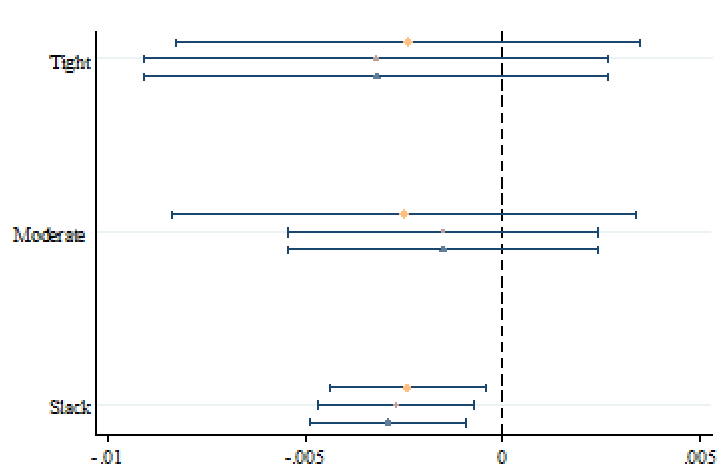


Figure B.14: Effects of OverWage on Target by Occupation Tightness

Figure B.14a: Coefficients of Overwage on Target

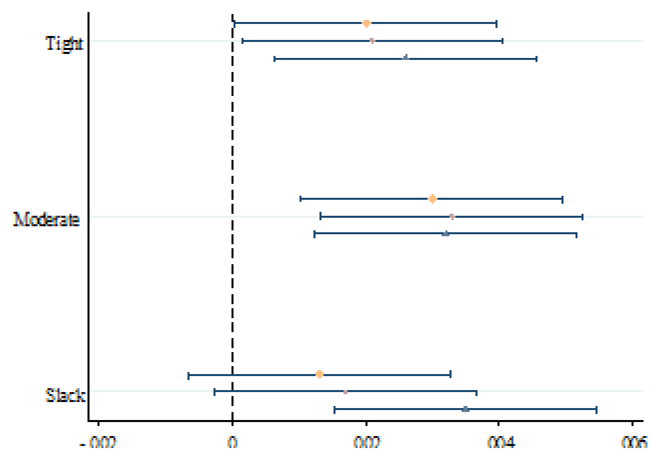



Figure B.14b: Coefficients of Underwage*Disclose on Target

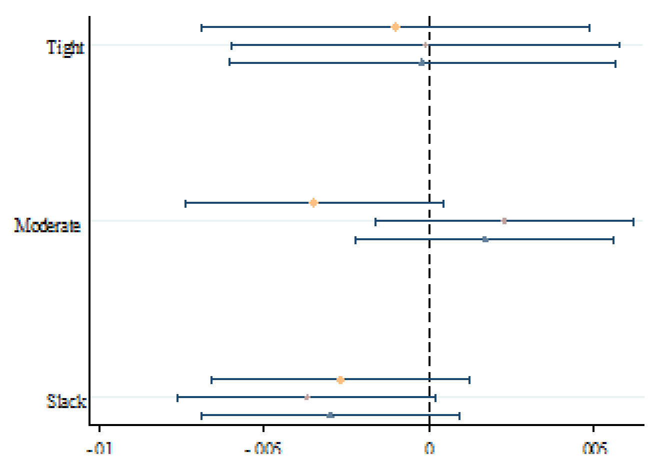
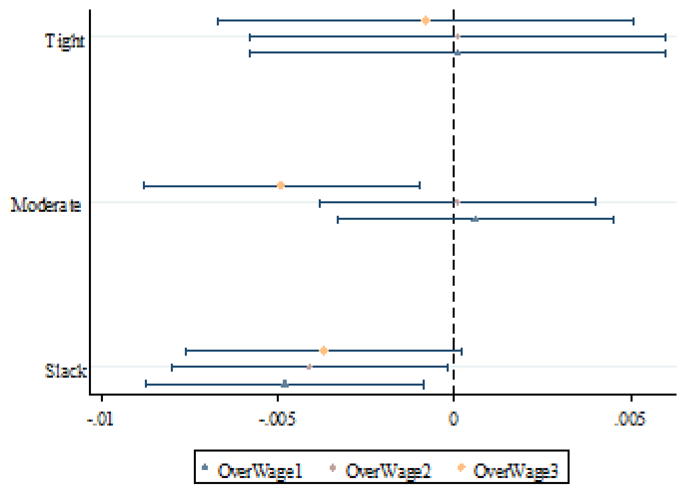


Figure B.14c: Coefficients of Overwage*Disclose on Target




Notes:

1. Tight industries include Advertising/Marketing, Finance, Automobiles/Manufacturing, and Government/Agriculture. Moderate industries include Transportation/Trade, Professional Services, Commodity, and Real Estate/Construction. Slack industries include Electronics/Computer Hardware, Internet/Computer Software, Energy and Medical/Pharmaceuticals.

2. Tight occupations include Traffic Service, Film Entertainment/Media, Official/Public Service/Science Research, Hospitality/Restaurant/Entertainment, Translator, Banking, Telecommunication/Communication Technology, Manufacturing/Operation, Tourism/Exit and Entry Service, Fund/Security/Futures/Investments, Project Management, Senior Management, Construction, Electricity/Energy/Mining, Automotive Sales and Service, Mechanical Design/Production/Maintenance, IT Management/Project Coordination, Apparels/Textiles/Leather Goods. Moderate occupations include Automobile Manufacture, Department Store/Chain Shops/Retail, Logistics/Warehouse, Accounting/Auditing/Taxation, Writing/Newspaper/Publishing/Printing, Purchasing/Trade, Environmental Science/Environmental, Advertising/Exhibition, Hospital/Medicine/Nursing, Consultant/Research, Electronics/Wiring/Semiconductor, Chemical, Marketing, Real Estate Development, Public Relations/Media, Trust/Guarantee/Auction/Pawn Business, Agriculture/Forestry/Animal Husbandry/F, Human Resource, Real Estate Agent/Broker. Slack occupations include Product/Operation/Design, IT QA/Testing/Configuration Management, Admin./Support Services/Secretarial, Education/Training, Sales Management, Legal/Compliance, Property Management, Intern/Trainee/Associate Trainee, Quality Management/Safety Protection, Biotechnology/Pharmaceuticals/Medical E, Art/Design, Insurance, IT Operation/Technical Support, Customer Service/Technical Support, Software Development/System Integration, Salespersons, Hardware Development, Sales Administration.

Table B.20: The Effect of Wage Disclosure on Becoming a Recruiting Target (OverWage2 in Viewed Sample)

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| Disclose | 0.0054*** | 0.0052*** | 0.0041*** | 0.0049*** | 0.0044*** | 0.0047*** | 0.0019*** |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.001) |
| OverWage2 | -0.0036*** | -0.0051*** | -0.0034*** | -0.0025*** | -0.0027*** | -0.0027*** | -0.0020 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.001) |
| Disclose* OverWage2 | -0.0031*** | -0.0046*** | -0.0034*** | -0.0025*** | -0.0027*** | -0.0026*** | -0.0026** |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.001) |
| Job Characteristics | | Yes | Yes | Yes | Yes | | |
| Job Match | | | Yes | Yes | Yes | Yes | Yes |
| Worker Characteristics | | | | Yes | Yes | Yes | |
| City, Time, Ind & Occ FE | | | | | Yes | | |
| Firm FE | | | | | Yes | | |
| Job FE | | | | | | Yes | Yes |
| Worker FE | | | | | | | Yes |
| ‘Effective’ *N* | 3,542,049 | 3,542,049 | 3,542,049 | 3,542,049 | 3,541,342 | 3,486,460 | 3,201,726 |

Standard errors in parentheses, clustered by job. *** p<0.01, ** p<0.05, * p<0.1

Note: In Section 4.2, Approach 2 derives the expectation on candidates’ wage as the median wage of all applicants. Here we use the revealed median wage as the expected wage, and define Overwage2 = 1 if the worker has a wage above the median wage of applicants who choose to disclose. The specifications are the sample with Panel B in Table 2.

### B.5.5 Heterogeneity of Employer's Response

Table B.21: The Effect of Gender on Becoming a Recruiting Target

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| **Panel A** | | | | | | | |
| OverWage1 | 0.0111*** | 0.0092*** | 0.0082*** | 0.0091*** | 0.0067*** | 0.0070*** | 0.0048*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.000) | (0.001) |
| Disclose*UnderWage1 | -0.0020*** | -0.0031*** | -0.0013* | -0.0010 | -0.0024*** | -0.0022*** | 0.0002 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.000) | (0.000) | (0.002) |
| Disclose*OverWage1 | -0.0028*** | -0.0042*** | -0.0027*** | -0.0024*** | -0.0030*** | -0.0035*** | -0.0018 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.002) |
| Male*UnderWage1 | 0.0005 | 0.0008 | 0.0019 | 0.0024*** | -0.0001 | -0.0001 | 0.0071 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.000) | (0.000) | (0.005) |
| Male*OverWage1 | -0.0017 | -0.0016 | -0.0006 | 0.0002 | -0.0017*** | -0.0021*** | 0.0047 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.000) | (0.000) | (0.005) |
| Male*Disclose* UnderWage1 | -0.0021*** | -0.0024*** | -0.0025*** | -0.0021** | -0.0010 | -0.0011** | -0.0021 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.002) |
| Male*Disclose*OverWage1 | 0.0000 | -0.0001 | -0.0004 | -0.0001 | 0.0001 | 0.0006 | -0.0009 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.002) |
| **Panel B** | | | | | | | |
| OverWage2 | 0.0061*** | 0.0056*** | 0.0045*** | 0.0057*** | 0.0047*** | 0.0053*** | 0.0037*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.000) | (0.000) | (0.001) |
| Disclose*UnderWage2 | -0.0025*** | -0.0039*** | -0.0021*** | -0.0018*** | -0.0026*** | -0.0023*** | 0.0000 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.000) | (0.000) | (0.002) |
| Disclose*OverWage2 | -0.0020*** | -0.0034*** | -0.0019*** | -0.0014** | -0.0027*** | -0.0033*** | -0.0016 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.000) | (0.000) | (0.002) |
| Male*UnderWage2 | -0.0003 | -0.0002 | 0.0010 | 0.0016** | -0.0008* | -0.0007* | 0.0072 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.000) | (0.000) | (0.005) |
| Male*OverWage2 | -0.0010 | -0.0005 | 0.0005 | 0.0010 | -0.0010** | -0.0014*** | 0.0049 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.000) | (0.000) | (0.005) |
| Male*Disclose*UnderWage2 | -0.0011 | -0.0014* | -0.0014* | -0.0010 | -0.0005 | -0.0009 | -0.0020 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.002) |
| Male*Disclose*OverWage2 | -0.0011 | -0.0013* | -0.0016** | -0.0013* | -0.0003 | 0.0004 | -0.0010 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.002) |
| **Panel C** | | | | | | | |
| OverWage3 | 0.0069*** | 0.0067*** | 0.0064*** | 0.0067*** | 0.0065*** | 0.0069*** | 0.0030*** |
| | (0.000) | (0.000) | (0.000) | (0.001) | (0.000) | (0.000) | (0.001) |
| Disclose*UnderWage3 | -0.0030*** | -0.0039*** | -0.0021*** | -0.0015** | -0.0021*** | -0.0019*** | 0.0001 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.000) | (0.000) | (0.002) |
| Disclose*OverWage3 | -0.0035*** | -0.0045*** | -0.0030*** | -0.0025*** | -0.0036*** | -0.0039*** | -0.0022 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.002) |
| Male*UnderWage3 | -0.0008 | -0.0008 | 0.0006 | 0.0012* | -0.0005 | -0.0006 | 0.0078 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.000) | (0.000) | (0.005) |
| Male*OverWage3 | -0.0000 | 0.0001 | 0.0011 | 0.0017** | -0.0003 | -0.0005 | 0.0059 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.000) | (0.005) |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Male*Disclose*UnderWage3 | -0.0009 | -0.0013* | -0.0015** | -0.0012 | -0.0008 | -0.0009 | -0.0026 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.002) |
| Male*Disclose*OverWage3 | -0.0000 | -0.0002 | -0.0004 | -0.0002 | 0.0006 | 0.0008 | -0.0003 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.002) |
| Job Characteristics | | Yes | Yes | Yes | Yes | | |
| Job Match | | | Yes | Yes | Yes | Yes | Yes |
| Worker Characteristics | | | | Yes | Yes | Yes | |
| City, Time, Ind & Occ FE | | | | | Yes | | |
| Firm FE | | | | | Yes | | |
| Job FE | | | | | | Yes | Yes |
| Worker FE | | | | | | | Yes |
| 'Effective' *N* | 3,542,049 | 3,542,049 | 3,542,049 | 3,542,049 | 3,541,342 | 3,486,460 | 3,201,726 |

Standard errors in parentheses, clustered by job. *** p<0.01, ** p<0.05, * p<0.1

### B.5.6 Success Rate in the Batch Sample

Table B.22: The Effect of Wage Disclosure on Becoming a Recruiting Target

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| **Panel A** | | | | | | | |
| OverWage1 | 0.0071*** | 0.0054*** | 0.0046*** | 0.0053*** | 0.0037*** | 0.0039*** | 0.0029*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.000) | (0.000) | (0.001) |
| Disclose* UnderWage1 | -0.0022*** | -0.0028*** | -0.0016*** | -0.0018*** | -0.0020*** | -0.0020*** | -0.0022 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.000) | (0.000) | (0.002) |
| Disclose* OverWage1 | -0.0013** | -0.0021*** | -0.0012** | -0.0014** | -0.0015*** | -0.0015*** | -0.0023 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.000) | (0.000) | (0.002) |
| **Panel B** | | | | | | | |
| OverWage2 | 0.0048*** | 0.0044*** | 0.0035*** | 0.0042*** | 0.0031*** | 0.0033*** | 0.0024*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.000) | (0.000) | (0.001) |
| Disclose* UnderWage2 | -0.0019*** | -0.0027*** | -0.0015** | -0.0017*** | -0.0022*** | -0.0023*** | -0.0023 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.000) | (0.000) | (0.002) |
| Disclose* OverWage2 | -0.0014** | -0.0022*** | -0.0014** | -0.0015*** | -0.0014*** | -0.0013*** | -0.0022 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.000) | (0.000) | (0.002) |
| **Panel C** | | | | | | | |
| OverWage3 | 0.0061*** | 0.0057*** | 0.0053*** | 0.0054*** | 0.0048*** | 0.0051*** | 0.0015** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.000) | (0.000) | (0.001) |
| Disclose* UnderWage3 | -0.0026*** | -0.0032*** | -0.0021*** | -0.0022*** | -0.0023*** | -0.0021*** | -0.0028 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.000) | (0.000) | (0.002) |
| Disclose* OverWage3 | -0.0011* | -0.0018*** | -0.0008 | -0.0010 | -0.0011** | -0.0012** | -0.0020 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.002) |
| Job Characteristics | | Yes | Yes | Yes | Yes | | |
| Job Match | | | Yes | Yes | Yes | Yes | Yes |
| Worker Characteristics | | | | Yes | Yes | Yes | |
| City, Time, Ind & Occ FE | | | | | Yes | | |
| Firm FE | | | | | Yes | | |
| Job FE | | | | | | Yes | Yes |
| Worker FE | | | | | | | Yes |
| 'Effective' *N* | 907,165 | 907,165 | 907,165 | 907,165 | 906,003 | 880,385 | 806,015 |

Standard errors in parentheses, clustered by job. *** p<0.01, ** p<0.05, * p<0.1

Note: This table runs the regressions of Table 2 in the batch application sample.

## B.6 Employer's Response on Wages of Successful Applications

### B.6.1 Extensions of Table 2.3

Table B.23: The Effect of Wage Disclosure on the Lower Posted Wage in Successful Applications

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| | **Panel A** | | | | | | |
| OverWage1 | 0.6655*** | 1.2961*** | 1.2996*** | 1.2875*** | 1.2850*** | 1.2599*** | 1.3047*** |
| | (0.169) | (0.218) | (0.214) | (0.212) | (0.212) | (0.210) | (0.186) |
| Disclose* UnderWage1 | -3.0674*** | -2.1542*** | -1.8795*** | -1.4862*** | -1.4593*** | -1.2258*** | -0.4616 |
| | (0.102) | (0.071) | (0.070) | (0.069) | (0.069) | (0.069) | (0.332) |
| Disclose* OverWage1 | -2.0878*** | 0.4605*** | 0.3577*** | 0.4430*** | 0.3517*** | 0.3117*** | 0.2800*** |
| | (0.044) | (0.050) | (0.049) | (0.048) | (0.048) | (0.046) | (0.108) |
| | **Panel B** | | | | | | |
| OverWage2 | 0.2828*** | 0.7747*** | 0.7759*** | 0.7713*** | 0.7691*** | 0.7465*** | 0.7800*** |
| | (0.065) | (0.167) | (0.161) | (0.158) | (0.157) | (0.153) | (0.145) |
| Disclose* UnderWage2 | -2.9134*** | -2.2547*** | -1.9853*** | -1.4519*** | -1.4180*** | -1.1841*** | -0.5199* |
| | (0.079) | (0.061) | (0.060) | (0.057) | (0.057) | (0.056) | (0.337) |
| Disclose* OverWage2 | -2.4339*** | -0.1061* | 0.0647 | 0.3973*** | 0.3092*** | 0.2730*** | 0.2933*** |
| | (0.069) | (0.058) | (0.057) | (0.058) | (0.058) | (0.056) | (0.102) |
| | **Panel C** | | | | | | |
| OverWage3 | 0.3208*** | 1.4362*** | 1.0583*** | 1.9317*** | 1.9071*** | 1.5799*** | 1.3787*** |
| | (0.071) | (0.151) | (0.145) | (0.141) | (0.140) | (0.140) | (0.133) |
| Disclose* UnderWage3 | -2.6339*** | -1.8527*** | -1.7370*** | -1.3184*** | -1.2875*** | -1.0970*** | -0.2424 |
| | (0.082) | (0.063) | (0.062) | (0.060) | (0.060) | (0.059) | (0.346) |
| Disclose* OverWage3 | -2.4898*** | 0.0010 | 0.0839 | 0.3526*** | 0.3569*** | 0.3952*** | 0.2169** |
| | (0.079) | (0.066) | (0.064) | (0.063) | (0.063) | (0.062) | (0.111) |
| Wage & Gender | | Yes | Yes | Yes | Yes | Yes | |
| Demographics | | | Yes | Yes | Yes | Yes | |
| Edu, Exp, Classification | | | | Yes | Yes | Yes | |
| Match & Application | | | | | Yes | Yes | |
| City, Time, Ind & Occ FE | | | | | | Yes | |
| Worker FE | | | | | | | Yes |
| 'Effective' *N* | 417,723 | 417,723 | 417,723 | 417,723 | 417,723 | 417,676 | 266,713 |

Standard errors in parentheses, clustered by worker's sub-occupation. *** p<0.01, ** p<0.05, * p<0.1

Note: This table has the same identifications in Table 3. The outcome variable is the lower bound of posted wage in successful applications.

Table B.24: The Effect of Wage Disclosure on the Upper Posted Wage in Successful Applications

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| | **Panel A** | | | | | | |
| OverWage1 | 0.8415*** | 1.7594*** | 1.7642*** | 1.7394*** | 1.7354*** | 1.6905*** | 1.7239*** |
| | (0.206) | (0.236) | (0.231) | (0.229) | (0.229) | (0.227) | (0.212) |
| Disclose* UnderWage1 | -4.8403*** | -3.5362*** | -3.1159*** | -2.3313*** | -2.3051*** | -1.7793*** | -0.3813 |
| | (0.140) | (0.099) | (0.097) | (0.093) | (0.093) | (0.089) | (0.305) |
| Disclose* OverWage1 | -3.5035*** | 0.1847*** | 0.4810*** | 0.4448*** | 0.3459*** | 0.2381*** | 0.2051* |
| | (0.073) | (0.062) | (0.062) | (0.061) | (0.061) | (0.060) | (0.110) |
| | **Panel B** | | | | | | |
| OverWage2 | 0.4278*** | 1.1447*** | 1.1502*** | 1.1460*** | 1.1427*** | 1.1104*** | 1.1997*** |
| | (0.092) | (0.192) | (0.186) | (0.182) | (0.181) | (0.175) | (0.179) |
| Disclose* UnderWage2 | -4.7465*** | -3.7932*** | -3.3678*** | -2.3547*** | -2.3184*** | -1.7865*** | -0.5931* |
| | (0.115) | (0.090) | (0.087) | (0.081) | (0.081) | (0.077) | (0.313) |
| Disclose* OverWage2 | -3.8698*** | -0.4445*** | -0.1842** | 0.4241*** | 0.4296*** | 0.3314*** | 0.2894*** |
| | (0.098) | (0.074) | (0.074) | (0.074) | (0.074) | (0.072) | (0.098) |
| | **Panel C** | | | | | | |
| OverWage3 | 0.4404*** | 1.4853*** | 1.0150*** | 1.8062*** | 1.7653*** | 1.1559*** | 1.1389*** |
| | (0.096) | (0.173) | (0.166) | (0.160) | (0.159) | (0.157) | (0.164) |
| Disclose* UnderWage3 | -4.2659*** | -3.1348*** | -2.9506*** | -2.1419*** | -2.1090*** | -1.6538*** | -0.1423 |
| | (0.117) | (0.091) | (0.090) | (0.084) | (0.084) | (0.080) | (0.328) |
| Disclose* OverWage3 | -3.9091*** | -0.2636*** | -0.1288 | 0.3837*** | 0.3759*** | 0.3346*** | 0.2930** |
| | (0.112) | (0.085) | (0.084) | (0.082) | (0.082) | (0.081) | (0.118) |
| Wage & Gender | | Yes | Yes | Yes | Yes | Yes | |
| Demographics | | | Yes | Yes | Yes | Yes | |
| Edu, Exp, Classification | | | | Yes | Yes | Yes | |
| Match & Application | | | | | Yes | Yes | |
| City, Time, Ind & Occ FE | | | | | | Yes | |
| Worker FE | | | | | | | Yes |
| 'Effective' *N* | 417,723 | 417,723 | 417,723 | 417,723 | 417,723 | 417,676 | 266,713 |

Standard errors in parentheses, clustered by worker's sub-occupation. *** p<0.01, ** p<0.05, * p<0.1

Note: This table has the same identifications in Table 3. The outcome variable is the lower bound of posted wage in successful applications.

Table B.25: The Effect of Wage Disclosure on the Posted Wage in Successful Applications (OverWage2 in Viewed Sample)

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| OverWage2 | 0.2969*** | 0.8406*** | 0.8711*** | 0.8017*** | 0.8762*** | 0.8284*** | 0.8813*** |
| | (0.093) | (0.156) | (0.149) | (0.145) | (0.144) | (0.141) | (0.174) |
| Disclose* UnderWage2 | -3.6396*** | -2.6756*** | -2.3182*** | -1.3962*** | -1.3514*** | -1.0804*** | -0.3728 |
| | (0.128) | (0.102) | (0.098) | (0.091) | (0.092) | (0.088) | (0.315) |
| Disclose* OverWage2 | -2.7042*** | -0.1872** | 0.0588 | 0.3439*** | 0.352*** | 0.2416*** | 0.2355** |
| | (0.105) | (0.077) | (0.074) | (0.073) | (0.072) | (0.071) | (0.106) |
| Wage & Gender | | Yes | Yes | Yes | Yes | Yes | |
| Demographics | | | Yes | Yes | Yes | Yes | |
| Edu, Exp, Classification | | | | Yes | Yes | Yes | |
| Match & Application | | | | | Yes | Yes | |
| City, Time, Ind & Occ FE | | | | | | Yes | |
| Worker FE | | | | | | | Yes |
| 'Effective' *N* | 417,723 | 417,723 | 417,723 | 417,723 | 417,723 | 417,676 | 266,713 |

Standard errors in parentheses, clustered by worker's sub-occupation. *** p<0.01, ** p<0.05, * p<0.1

Note: Note: In Section 4.2, Approach 2 derives the expectation on candidates' wage as the median wage of all applicants. Here we use the revealed median wage as the expected wage, and define Overwage2 = 1 if the worker has a wage above the median wage of applicants who choose to disclose. The specifications are the sample with Panel B in Table 3.

### B.6.2 Complete Version of Table 2.3

Table B.26: The Effect of Wage Disclosure on the Posted Wage in Successful Applications (Panel A)

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| OverWage1 | 0.7536*** | 1.5278*** | 1.5319*** | 1.5135*** | 1.5102*** | 1.4752*** | 1.3143*** |
| | (0.186) | (0.224) | (0.220) | (0.218) | (0.218) | (0.216) | (0.196) |
| Disclose* UnderWage1 | -3.9539*** | -2.8452*** | -2.4977*** | -1.9087*** | -1.8822*** | -1.5025*** | -0.4215 |
| | (0.120) | (0.083) | (0.082) | (0.079) | (0.079) | (0.077) | (0.261) |
| Disclose* OverWage1 | -2.7956*** | 0.3226*** | 0.5693*** | 0.4939*** | 0.4988*** | 0.3249*** | 0.2925** |
| | (0.058) | (0.054) | (0.054) | (0.053) | (0.053) | (0.052) | (0.136) |
| Worker wage | | 0.7602*** | 0.7055*** | 0.6632*** | 0.6624*** | 0.5893*** | |
| | | (0.010) | (0.010) | (0.010) | (0.010) | (0.010) | |
| Male | | 1.4857*** | 0.8868*** | 0.8234*** | 0.8223*** | 0.8498*** | |
| | | (0.062) | (0.060) | (0.056) | (0.056) | (0.049) | |
| Age | | | 1.2152*** | 0.0961 | 0.0830 | 0.2551*** | |
| | | | (0.087) | (0.094) | (0.094) | (0.086) | |
| Age^2 | | | -0.0128*** | 0.0012 | 0.0014 | -0.0008 | |
| | | | (0.001) | (0.001) | (0.001) | (0.001) | |
| Single | | | 1.0103*** | 0.8507*** | 0.9541*** | 0.9389*** | |
| | | | (0.142) | (0.140) | (0.119) | (0.112) | |
| Married | | | 0.3949*** | 0.1826*** | 0.1985*** | 0.3172*** | |
| | | | (0.064) | (0.062) | (0.061) | (0.059) | |
| Worker experience | | | | 0.2909*** | 0.2898*** | 0.2993*** | |
| | | | | (0.033) | (0.033) | (0.033) | |
| Worker experience^2 | | | | -0.0067*** | -0.0067*** | -0.0071*** | |
| | | | | (0.001) | (0.001) | (0.002) | |
| Post doc | | | | 6.2594*** | 6.2473*** | 6.4114*** | |
| | | | | (1.495) | (1.494) | (1.531) | |
| Phd | | | | 4.8296*** | 4.8133*** | 4.5209*** | |
| | | | | (0.760) | (0.759) | (0.755) | |
| MBA | | | | 2.3303*** | 2.3414*** | 1.6685*** | |
| | | | | (0.560) | (0.559) | (0.550) | |
| Master | | | | 1.9186*** | 1.9069*** | 1.2330** | |
| | | | | (0.518) | (0.518) | (0.520) | |
| Bachelor | | | | 0.4739 | 0.4710 | 0.1201 | |
| | | | | (0.515) | (0.515) | (0.520) | |
| Tech college | | | | -0.7841 | -0.7771 | -0.9286* | |
| | | | | (0.515) | (0.514) | (0.520) | |
| Secondary | | | | -0.9534* | -0.9675* | -0.9151* | |
| | | | | (0.556) | (0.556) | (0.552) | |
| Tongzhao | | | | 0.5369*** | 0.5280*** | 0.6144*** | |
| | | | | (0.065) | (0.065) | (0.064) | |
| 985/211 | | | | 0.2757*** | 0.2708*** | 0.2733*** | |
| | | | | (0.090) | (0.090) | (0.088) | |
| Intensively Search | | | | 0.4056*** | 0.4241*** | 0.5725*** | |
| | | | | (0.051) | (0.051) | (0.050) | |
| Moderately Search | | | | 2.5845*** | 2.5359** | 2.1926** | |

| | | | |
|---|---|---|---|
| | (0.983) | (0.985) | (0.966) |
| Stay | 1.1475*** | 1.1460*** | 1.3351*** |
| | (0.047) | (0.047) | (0.045) |
| Tenure 1 | -0.0676*** | -0.0688*** | -0.0143 |
| | (0.012) | (0.012) | (0.012) |
| Tenure 2 | -0.0901*** | -0.0890*** | -0.0181 |
| | (0.013) | (0.013) | (0.013) |
| Domes 1-19 | 0.5792*** | 0.5691*** | 0.4617*** |
| | (0.114) | (0.114) | (0.114) |
| Domes 20-39 | 0.2515** | 0.2569** | 0.2211* |
| | (0.122) | (0.121) | (0.117) |
| Domes 40-59 | -0.2015 | -0.1956 | -0.1653 |
| | (0.123) | (0.123) | (0.122) |
| Domes 60-99 | 0.2891*** | 0.2872*** | 0.0512 |
| | (0.099) | (0.099) | (0.099) |
| Dome 100-199 | 0.0225 | 0.0390 | 0.1317** |
| | (0.057) | (0.057) | (0.056) |
| World 1-19 | 1.5763** | 1.5816** | 1.2287 |
| | (0.802) | (0.805) | (0.762) |
| World 20-39 | -0.9900*** | -0.9713*** | -0.5491* |
| | (0.329) | (0.329) | (0.323) |
| World 40-59 | 1.6619*** | 1.6233*** | 0.9830*** |
| | (0.286) | (0.286) | (0.277) |
| World 60-99 | 1.0247*** | 0.9828*** | 0.7803*** |
| | (0.304) | (0.304) | (0.290) |
| World 100-199 | 0.0820 | 0.0951 | 0.0203 |
| | (0.151) | (0.151) | (0.142) |
| Account days | 0.0704*** | 0.0699*** | 0.0617*** |
| | (0.004) | (0.004) | (0.004) |
| Resume Completeness | 0.4350*** | 0.4447*** | 0.2893*** |
| | (0.035) | (0.035) | (0.034) |
| Membership | 0.2400*** | 0.2477*** | 0.1361** |
| | (0.064) | (0.063) | (0.062) |
| Elite | 3.7453*** | 3.7306*** | 3.0660*** |
| | (0.078) | (0.078) | (0.077) |
| Batch Apply | | -0.0391 | -0.2539*** |
| | | (0.073) | (0.073) |
| Lens*1 | | -0.4317*** | -0.1001 |
| | | (0.084) | (0.082) |
| Lens*2 | | 0.0765 | 0.2103 |
| | | (0.137) | (0.135) |
| Log (wage gap) | | 0.0098*** | 0.0086*** |
| | | (0.003) | (0.002) |
| City Match | | 0.6631*** | -0.3943*** |
| | | (0.096) | (0.092) |
| Province Match | | -0.3730*** | -0.2342** |
| | | (0.105) | (0.100) |
| Main Industry Match | | 0.4683*** | -0.0439 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | (0.073) | (0.067) | |
| Sub-Industry Match | | | | | 0.0481 | 0.0220 | |
| | | | | | (0.062) | (0.059) | |
| Main Occupation Match | | | | | -0.1889** | -0.0434 | |
| | | | | | (0.074) | (0.070) | |
| Sub-Occupation Match | | | | | -0.0668 | -0.2143*** | |
| | | | | | (0.061) | (0.060) | |
| | | | | | | | |
| Time FE | | | | | | Yes | Yes |
| Worker Location Ind, and Occ FE | | | | | | Yes | |
| Worker FE | | | | | | | Yes |
| | | | | | | | |
| 'Effective' *N* | 417,723 | 417,723 | 417,723 | 417,723 | 417,723 | 417,676 | 266,713 |
| $R^2$ | 0.051 | 0.479 | 0.489 | 0.499 | 0.500 | 0.518 | 0.745 |

Standard errors in parentheses, clustered by job. *** p<0.01, ** p<0.05, * p<0.1

Table B.27: The Effect of Wage Disclosure on the Posted Wage in Successful Applications (Panel B)

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| OverWage2 | 0.3553*** | 0.9597*** | 0.9630*** | 0.9587*** | 0.9559*** | 0.9284*** | 0.9899*** |
| | (0.077) | (0.177) | (0.171) | (0.168) | (0.167) | (0.162) | (0.159) |
| Disclose* UnderWage2 | -3.8299*** | -3.0240*** | -2.6766*** | -1.9033*** | -1.8682*** | -1.4853*** | -0.4565* |
| | (0.096) | (0.074) | (0.072) | (0.067) | (0.067) | (0.065) | (0.268) |
| Disclose* OverWage2 | -3.1518*** | -0.2753*** | -0.0597 | 0.4107*** | 0.4194*** | 0.4522*** | 0.2914*** |
| | (0.082) | (0.065) | (0.064) | (0.065) | (0.064) | (0.063) | (0.101) |
| Worker wage | | 0.7331*** | 0.6804*** | 0.6340*** | 0.6328*** | 0.5541*** | |
| | | (0.011) | (0.011) | (0.011) | (0.011) | (0.011) | |
| Male | | 1.8535*** | 1.2711*** | 1.2284*** | 1.1949*** | 1.0922*** | |
| | | (0.081) | (0.079) | (0.074) | (0.072) | (0.056) | |
| Age | | | 1.0201*** | -0.0889 | -0.1018 | 0.1222 | |
| | | | (0.096) | (0.103) | (0.103) | (0.094) | |
| Age^2 | | | -0.0100*** | 0.0039** | 0.0041** | 0.0011 | |
| | | | (0.001) | (0.002) | (0.002) | (0.001) | |
| Single | | | 1.7973*** | 1.5406*** | 1.6453*** | 1.5107*** | |
| | | | (0.167) | (0.165) | (0.139) | (0.131) | |
| Married | | | 0.3249*** | 0.1220* | 0.1402** | 0.2850*** | |
| | | | (0.072) | (0.069) | (0.068) | (0.065) | |
| Worker experience | | | | 0.2666*** | 0.2697*** | 0.2982*** | |
| | | | | (0.034) | (0.035) | (0.034) | |
| Worker experience^2 | | | | -0.0057*** | -0.0058*** | -0.0068*** | |
| | | | | (0.002) | (0.002) | (0.002) | |
| Post doc | | | | 9.1548*** | 9.1810*** | 8.9292*** | |
| | | | | (1.641) | (1.642) | (1.675) | |
| Phd | | | | 7.6327*** | 7.6461*** | 6.9039*** | |
| | | | | (0.821) | (0.819) | (0.809) | |
| MBA | | | | 4.5041*** | 4.5495*** | 3.4972*** | |
| | | | | (0.596) | (0.595) | (0.577) | |
| Master | | | | 3.9606*** | 3.9789*** | 2.8979*** | |
| | | | | (0.535) | (0.535) | (0.532) | |
| Bachelor | | | | 2.1510*** | 2.1758*** | 1.4939*** | |
| | | | | (0.531) | (0.531) | (0.530) | |
| Tech college | | | | 0.2800 | 0.3084 | -0.0961 | |
| | | | | (0.530) | (0.530) | (0.531) | |
| Secondary | | | | -0.4639 | -0.4764 | -0.5034 | |
| | | | | (0.566) | (0.565) | (0.558) | |
| Tongzhao | | | | 0.7572*** | 0.7510*** | 0.8427*** | |
| | | | | (0.072) | (0.072) | (0.070) | |
| 985/211 | | | | 0.2358** | 0.2330** | 0.2608*** | |
| | | | | (0.098) | (0.098) | (0.096) | |
| Intensively Search | | | | 0.6438*** | 0.6580*** | 0.8391*** | |
| | | | | (0.057) | (0.057) | (0.054) | |
| Moderately Search | | | | 3.7246*** | 3.6614*** | 3.1450*** | |

| | | | |
|---|---|---|---|
| | (1.097) | (1.099) | (1.084) |
| Stay | 1.5590*** | 1.5520*** | 1.7604*** |
| | (0.053) | (0.053) | (0.051) |
| Tenure 1 | -0.0624*** | -0.0648*** | -0.0019 |
| | (0.013) | (0.013) | (0.013) |
| Tenure 2 | -0.0981*** | -0.0978*** | -0.0198 |
| | (0.014) | (0.014) | (0.014) |
| Domes 1-19 | 0.7314*** | 0.7221*** | 0.5992*** |
| | (0.128) | (0.127) | (0.126) |
| Domes 20-39 | 0.3074** | 0.3141** | 0.2854** |
| | (0.135) | (0.134) | (0.127) |
| Domes 40-59 | -0.0863 | -0.0786 | -0.0171 |
| | (0.135) | (0.135) | (0.133) |
| Domes 60-99 | 0.3584*** | 0.3664*** | 0.1399 |
| | (0.109) | (0.109) | (0.108) |
| Dome 100-199 | 0.1334** | 0.1493** | 0.2106*** |
| | (0.064) | (0.064) | (0.062) |
| World 1-19 | 1.8119** | 1.7842** | 1.3492 |
| | (0.865) | (0.866) | (0.827) |
| World 20-39 | -0.6894* | -0.6893* | -0.3599 |
| | (0.379) | (0.378) | (0.366) |
| World 40-59 | 1.8984*** | 1.8608*** | 1.1172*** |
| | (0.319) | (0.319) | (0.307) |
| World 60-99 | 0.9680*** | 0.9274*** | 0.9013*** |
| | (0.338) | (0.338) | (0.322) |
| World 100-199 | -0.2033 | -0.1856 | -0.1113 |
| | (0.172) | (0.172) | (0.158) |
| Account days | 0.0698*** | 0.0702*** | 0.0640*** |
| | (0.005) | (0.005) | (0.005) |
| Resume Completeness | 0.3968*** | 0.3986*** | 0.2636*** |
| | (0.039) | (0.039) | (0.038) |
| Membership | 0.2172*** | 0.2242*** | 0.1181* |
| | (0.070) | (0.070) | (0.068) |
| Elite | 4.2103*** | 4.1608*** | 3.3131*** |
| | (0.090) | (0.092) | (0.098) |
| Batch Apply | | 0.0966 | -0.1528* |
| | | (0.083) | (0.082) |
| Lens*1 | | -0.3737*** | -0.0308 |
| | | (0.094) | (0.091) |
| Lens*2 | | 0.1689 | 0.3176** |
| | | (0.152) | (0.150) |
| Log (wage gap) | | 0.0114*** | 0.0100*** |
| | | (0.003) | (0.003) |
| City Match | | 0.6575*** | -0.4207*** |
| | | (0.110) | (0.102) |
| Province Match | | -0.5104*** | -0.3007*** |
| | | (0.120) | (0.111) |
| Main Industry Match | | 0.4902*** | -0.1302* |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | (0.083) | (0.076) | |
| Sub-Industry Match | | | | | 0.2594*** | 0.1689** | |
| | | | | | (0.069) | (0.066) | |
| Main Occupation Match | | | | | -0.1408 | -0.0697 | |
| | | | | | (0.087) | (0.080) | |
| Sub-Occupation Match | | | | | -0.1372** | -0.3098*** | |
| | | | | | (0.069) | (0.065) | |
| | | | | | | | |
| Time FE | | | | | | Yes | Yes |
| Worker Location Ind, and Occ FE | | | | | | Yes | |
| Worker FE | | | | | | | Yes |
| | | | | | | | |
| 'Effective' *N* | 417,723 | 417,723 | 417,723 | 417,723 | 417,723 | 417,676 | 266,713 |
| $R^2$ | 0.010 | 0.380 | 0.390 | 0.404 | 0.406 | 0.428 | 0.698 |

Standard errors in parentheses, clustered by job. *** p<0.01, ** p<0.05, * p<0.1

Table B.28: The Effect of Wage Disclosure on the Posted Wage in Successful Applications (Panel C)

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| OverWage3 | 0.3806*** | 1.9608*** | 1.5366*** | 1.3689*** | 1.3362*** | 1.8679*** | 1.2588*** |
| | (0.082) | (0.160) | (0.154) | (0.149) | (0.148) | (0.147) | (0.146) |
| Disclose* UnderWage3 | -3.4499*** | -2.4938*** | -2.3438*** | -1.7302*** | -1.6982*** | -1.3754*** | -0.1924 |
| | (0.098) | (0.075) | (0.074) | (0.070) | (0.071) | (0.068) | (0.279) |
| Disclose* OverWage3 | -3.1995*** | -0.1313* | -0.0225 | 0.3682*** | 0.3664*** | 0.3649*** | 0.2049** |
| | (0.094) | (0.073) | (0.072) | (0.071) | (0.071) | (0.070) | (0.100) |
| Worker wage | | 0.7079*** | 0.6688*** | 0.6286*** | 0.6273*** | 0.5458*** | |
| | | (0.012) | (0.012) | (0.012) | (0.012) | (0.012) | |
| Male | | 1.6208*** | 1.2660*** | 1.1320*** | 1.0793*** | 0.9060*** | |
| | | (0.086) | (0.084) | (0.080) | (0.078) | (0.063) | |
| Age | | | 0.4032*** | -0.4347*** | -0.4409*** | -0.2277** | |
| | | | (0.115) | (0.122) | (0.121) | (0.111) | |
| Age^2 | | | -0.0027 | 0.0080*** | 0.0081*** | 0.0050*** | |
| | | | (0.002) | (0.002) | (0.002) | (0.002) | |
| Single | | | 1.6951*** | 1.4609*** | 1.5346*** | 1.4076*** | |
| | | | (0.184) | (0.182) | (0.153) | (0.143) | |
| Married | | | 0.1942** | 0.0168 | 0.0340 | 0.1563** | |
| | | | (0.077) | (0.074) | (0.072) | (0.068) | |
| Worker experience | | | | 0.1655*** | 0.1695*** | 0.1962*** | |
| | | | | (0.040) | (0.040) | (0.039) | |
| Worker experience^2 | | | | -0.0033* | -0.0034* | -0.0045** | |
| | | | | (0.002) | (0.002) | (0.002) | |
| Post doc | | | | - | - | - | |
| Phd | | | | 6.2014*** | 6.2554*** | 5.5029*** | |
| | | | | (0.813) | (0.810) | (0.784) | |
| MBA | | | | 3.2228*** | 3.3037*** | 2.2096*** | |
| | | | | (0.531) | (0.527) | (0.492) | |
| Master | | | | 2.4848*** | 2.5418*** | 1.5003*** | |
| | | | | (0.439) | (0.436) | (0.417) | |
| Bachelor | | | | 1.4848*** | 1.5468*** | 0.8865** | |
| | | | | (0.425) | (0.423) | (0.402) | |
| Tech college | | | | 0.0020 | 0.0592 | -0.3708 | |
| | | | | (0.421) | (0.418) | (0.395) | |
| Secondary | | | | -0.5679 | -0.5661 | -0.6096 | |
| | | | | (0.494) | (0.491) | (0.466) | |
| Tongzhao | | | | 0.8262*** | 0.8236*** | 0.9350*** | |
| | | | | (0.075) | (0.075) | (0.073) | |
| 985/211 | | | | 0.2122** | 0.2163** | 0.2631*** | |
| | | | | (0.104) | (0.104) | (0.101) | |
| Intensively Search | | | | 0.8333*** | 0.8409*** | 1.0058*** | |
| | | | | (0.062) | (0.062) | (0.059) | |
| Moderately Search | | | | 3.8411*** | 3.7754*** | 3.3350*** | |

| | | | |
|---|---|---|---|
| | (1.215) | (1.216) | (1.212) |
| Stay | 1.8179*** | 1.8081*** | 1.9775*** |
| | (0.059) | (0.059) | (0.057) |
| Tenure 1 | -0.0678*** | -0.0708*** | -0.0064 |
| | (0.014) | (0.014) | (0.014) |
| Tenure 2 | -0.0965*** | -0.0969*** | -0.0234 |
| | (0.016) | (0.016) | (0.016) |
| Domes 1-19 | 0.7417*** | 0.7290*** | 0.6047*** |
| | (0.137) | (0.137) | (0.136) |
| Domes 20-39 | 0.3719** | 0.3725** | 0.3393** |
| | (0.146) | (0.145) | (0.138) |
| Domes 40-59 | -0.0915 | -0.0876 | -0.0527 |
| | (0.145) | (0.145) | (0.142) |
| Domes 60-99 | 0.3631*** | 0.3726*** | 0.1421 |
| | (0.119) | (0.119) | (0.118) |
| Dome 100-199 | 0.2379*** | 0.2499*** | 0.2689*** |
| | (0.068) | (0.068) | (0.065) |
| World 1-19 | 1.6843* | 1.6324* | 1.2112 |
| | (0.920) | (0.920) | (0.875) |
| World 20-39 | -0.6642 | -0.6609 | -0.1137 |
| | (0.413) | (0.412) | (0.399) |
| World 40-59 | 2.0541*** | 2.0222*** | 1.2639*** |
| | (0.333) | (0.333) | (0.320) |
| World 60-99 | 1.0001*** | 0.9762*** | 1.0551*** |
| | (0.358) | (0.357) | (0.339) |
| World 100-199 | -0.3888** | -0.3801** | -0.2672 |
| | (0.180) | (0.179) | (0.163) |
| Account days | 0.0721*** | 0.0723*** | 0.0667*** |
| | (0.005) | (0.005) | (0.005) |
| Resume Completeness | 0.3639*** | 0.3557*** | 0.2089*** |
| | (0.042) | (0.042) | (0.040) |
| Membership | 0.2587*** | 0.2593*** | 0.1510** |
| | (0.073) | (0.073) | (0.071) |
| Elite | 4.0192*** | 3.9550*** | 3.0931*** |
| | (0.099) | (0.100) | (0.111) |
| Batch Apply | | 0.2664*** | 0.0465 |
| | | (0.091) | (0.090) |
| Lens*1 | | -0.2292** | 0.0839 |
| | | (0.104) | (0.101) |
| Lens*2 | | 0.1532 | 0.2946* |
| | | (0.171) | (0.169) |
| Log (wage gap) | | 0.0109*** | 0.0097*** |
| | | (0.003) | (0.003) |
| City Match | | 0.3486*** | -0.5437*** |
| | | (0.126) | (0.116) |
| Province Match | | -0.2360* | -0.0925 |
| | | (0.137) | (0.126) |
| Main Industry Match | | 0.4671*** | -0.0986 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | (0.091) | (0.084) | |
| Sub-Industry Match | | | | | 0.2796*** | 0.1767** | |
| | | | | | (0.075) | (0.070) | |
| Main Occupation Match | | | | | -0.1389 | -0.0323 | |
| | | | | | (0.093) | (0.089) | |
| Sub-Occupation Match | | | | | -0.1768** | -0.3632*** | |
| | | | | | (0.072) | (0.070) | |
| | | | | | | | |
| Time FE | | | | | | Yes | Yes |
| Worker Location Ind, and Occ FE | | | | | | Yes | |
| Worker FE | | | | | | | Yes |
| | | | | | | | |
| ‘Effective’ *N* | 368,962 | 368,962 | 368,962 | 368,962 | 368,962 | 368,914 | 229,450 |
| $R^2$ | 0.009 | 0.376 | 0.380 | 0.392 | 0.393 | 0.416 | 0.697 |

Standard errors in parentheses, clustered by job. *** p<0.01, ** p<0.05, * p<0.1

## B.7 Discussions on Marital Status Revealing

### B.7.1 Worker's Characteristics and Marital Status Revealing

When workers set up their profiles, they are required to select their marital status from three options: 1) Single, 2) Married, 3) Confidential. Different from the wage revealing setting, where we can observe both workers' wages as well as their disclosure option, if a worker chooses his marital status as "confidential", the platform does not know his actual marital status.

According to the results from Table 2.1, there is a negative relationship between revealing wage and revealing marital status. To better describe the characteristics of workers that contribute to the marital status revealing, we regress the disclosure of marital status on a set of worker's characteristics used in Table 2.1 (except the high-wage worker indicators), and report the estimation results in Table B.29. The independent variable is 1 if the worker discloses his marital status (chooses his marital status as "Single" OR "Married").

Consistent with the results in wage revealing, male, especially older male workers are more likely to disclose their marital status. In addition, the increase of wage, years of working experience and education level leads to higher probability of marital status disclosure. Workers who disclose their current wages are unlikely to disclose marital status.

Table B.29: The Effect of Applicants' Characteristics on Disclosing Marital Status

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| Male | 0.0757*** | 0.0397*** | 0.0428*** | 0.0393*** | 0.0345*** | 0.0249*** | 0.0281*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Age | | 0.0202*** | 0.0158*** | 0.0156*** | 0.0152*** | 0.0145*** | 0.0136*** |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Age^2 | | -0.0008*** | -0.0003*** | -0.0003*** | -0.0003*** | -0.0003*** | -0.0003*** |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Age*Male | | 0.0040*** | 0.0036*** | 0.0021*** | 0.0019*** | 0.0026*** | 0.0031*** |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Age^2*Male | | -0.0001*** | -0.0001*** | -0.0000 | -0.0000 | -0.0000** | -0.0000 |
| | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Wage | | | 0.0005*** | 0.0003*** | 0.0006*** | 0.0007*** | 0.0008*** |
| | | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Disclose Wage | | | -0.0188*** | -0.0174*** | -0.0192*** | -0.0163*** | -0.0155*** |
| | | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Experience | | | | 0.0065*** | 0.0062*** | 0.0060*** | 0.0055*** |
| | | | | (0.000) | (0.000) | (0.000) | (0.000) |
| Experience^2 | | | | -0.0003*** | -0.0003*** | -0.0003*** | -0.0003*** |
| | | | | (0.000) | (0.000) | (0.000) | (0.000) |
| Post doc | | | | 0.0470** | 0.0477** | 0.0704*** | 0.0747*** |
| | | | | (0.020) | (0.020) | (0.020) | (0.020) |
| Phd | | | | 0.0443*** | 0.0465*** | 0.0695*** | 0.0730*** |
| | | | | (0.008) | (0.009) | (0.009) | (0.009) |
| MBA | | | | 0.0590*** | 0.0570*** | 0.0650*** | 0.0630*** |
| | | | | (0.006) | (0.006) | (0.006) | (0.006) |
| Master | | | | 0.0521*** | 0.0529*** | 0.0658*** | 0.0637*** |
| | | | | (0.006) | (0.006) | (0.006) | (0.006) |
| Bachelor | | | | 0.0610*** | 0.0597*** | 0.0600*** | 0.0552*** |
| | | | | (0.005) | (0.005) | (0.005) | (0.005) |
| Tech college | | | | 0.0421*** | 0.0401*** | 0.0375*** | 0.0349*** |
| | | | | (0.005) | (0.005) | (0.005) | (0.005) |
| Secondary | | | | 0.0060 | 0.0044 | 0.0028 | 0.0047 |
| | | | | (0.007) | (0.007) | (0.007) | (0.007) |
| Tongzhao | | | | 0.0013 | 0.0012 | 0.0007 | 0.0027** |
| | | | | (0.001) | (0.001) | (0.001) | (0.001) |
| 985/211 | | | | -0.0042** | -0.0025 | -0.0016 | -0.0009 |
| | | | | (0.002) | (0.002) | (0.002) | (0.002) |
| Intensively Search | | | | 0.0189*** | 0.0196*** | 0.0181*** | 0.0164*** |
| | | | | (0.001) | (0.001) | (0.001) | (0.001) |
| Moderately Search | | | | -0.0122 | -0.0104 | -0.0118 | -0.0155 |
| | | | | (0.019) | (0.019) | (0.019) | (0.019) |
| Stay | | | | 0.0112*** | 0.0109*** | 0.0114*** | 0.0096*** |
| | | | | (0.001) | (0.001) | (0.001) | (0.001) |
| Tenure 1 | | | | -0.0007*** | -0.0008*** | -0.0010*** | -0.0009*** |

| | | | | |
|---|---|---|---|---|
| | (0.000) | (0.000) | (0.000) | (0.000) |
| Tenure 2 | 0.0042*** | 0.0039*** | 0.0034*** | 0.0032*** |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| Domes 1-19 | 0.0152*** | 0.0162*** | 0.0147*** | 0.0121*** |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| Domes 20-39 | 0.0241*** | 0.0231*** | 0.0167*** | 0.0145*** |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| Domes 40-59 | 0.0163*** | 0.0159*** | 0.0117*** | 0.0091*** |
| | (0.003) | (0.003) | (0.003) | (0.003) |
| Domes 60-99 | 0.0118*** | 0.0133*** | 0.0143*** | 0.0124*** |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| Dome 100-199 | 0.0112*** | 0.0091*** | 0.0082*** | 0.0074*** |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| World 1-19 | -0.0473*** | -0.0461*** | -0.0239 | -0.0202 |
| | (0.016) | (0.016) | (0.016) | (0.016) |
| World 20-39 | 0.0070 | -0.0058 | 0.0036 | 0.0061 |
| | (0.009) | (0.009) | (0.009) | (0.009) |
| World 40-59 | -0.0251*** | -0.0262*** | -0.0170*** | -0.0143*** |
| | (0.005) | (0.005) | (0.005) | (0.005) |
| World 60-99 | -0.0336*** | -0.0221*** | -0.0124** | -0.0105* |
| | (0.006) | (0.006) | (0.006) | (0.006) |
| World 100-199 | -0.0123*** | -0.0068** | -0.0035 | -0.0018 |
| | (0.003) | (0.003) | (0.003) | (0.003) |
| Elite | | 0.0112*** | 0.0090*** | 0.0062*** |
| | | (0.002) | (0.002) | (0.002) |
| Account days | | 0.6352*** | 0.6292*** | 0.6197*** |
| | | (0.008) | (0.008) | (0.008) |
| Resume Completeness | | 1.7141*** | 1.5361*** | 1.4868*** |
| | | (0.089) | (0.088) | (0.089) |
| Membership | | -0.0172*** | -0.0158*** | -0.0158*** |
| | | (0.001) | (0.001) | (0.001) |
| Number of Applications | | -0.0002*** | -0.0002*** | -0.0002*** |
| | | (0.000) | (0.000) | (0.000) |
| Log (wage gap) | | 0.0112*** | 0.0090*** | 0.0062*** |
| | | (0.002) | (0.002) | (0.002) |
| City Match | | -0.0081*** | -0.0080*** | -0.0076*** |
| | | (0.002) | (0.002) | (0.002) |
| Province Match | | 0.0278*** | 0.0287*** | 0.0271*** |
| | | (0.002) | (0.002) | (0.002) |
| Main Industry Match | | -0.2574*** | -0.2745*** | -0.2675*** |
| | | (0.001) | (0.002) | (0.002) |
| Sub-Industry Match | | 0.1669*** | 0.1730*** | 0.1726*** |
| | | (0.001) | (0.001) | (0.001) |
| Main Occupation Match | | -0.1268*** | -0.1225*** | -0.1343*** |
| | | (0.002) | (0.002) | (0.002) |
| Sub-Occupation Match | | -0.0263*** | -0.0289*** | -0.0285*** |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | (0.002) | (0.002) | (0.002) |
| 'Effective' $N$ | 941,733 | 941,733 | 941,733 | 941,733 | 941,733 | 941,697 | 941,695 |
| $R^2$ | 0.006 | 0.058 | 0.063 | 0.134 | 0.136 | 0.144 | 0.148 |

Standard errors in parentheses, clustered by worker's sub-occupation. *** p<0.01, ** p<0.05, * p<0.1

Note: In this table, column 6 controls for the worker's location fixed effect, and column 7 further controls for the worker's industry and occupation fixed effects.

### B.7.2 Age and Marital Status Revealing

We can expect that people tend to disclose their marital status with the increase of age. In this section, we use worker's age as a proxy for the marital status, and predict the probability of marital status disclosure for both male and female workers with respect to age. The variables used in the prediction are age quartics and the interactions with other worker's characteristics, including gender, education, working experience, employment status, industry and occupation.

Figure B.15 to B.17 plot the predicted probability of marital status revealing and the difference between male and female workers. It suggests that males are more likely to disclose their marital status than females at any age, especially after 30. The gender difference in marital status revealing is mainly driven by the difference in disclosing as "married" and disclosing as "single" is not significantly different between male and female workers.

Figure B.15: Age and Marital Status Revealing

Figure B.15a: Share of Revealing Martial Status as a Function of Age

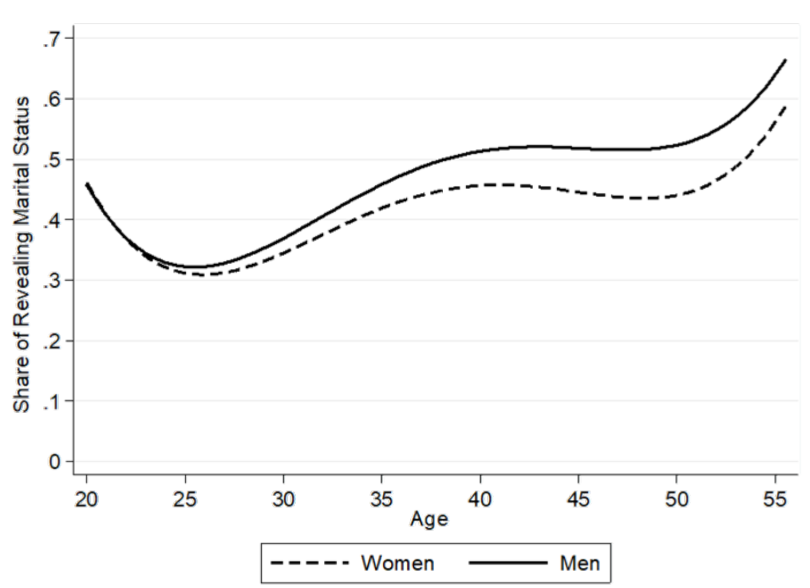



Figure B.15b: Predicted Male Share - Predicted Female Share in Revealing Martial Status

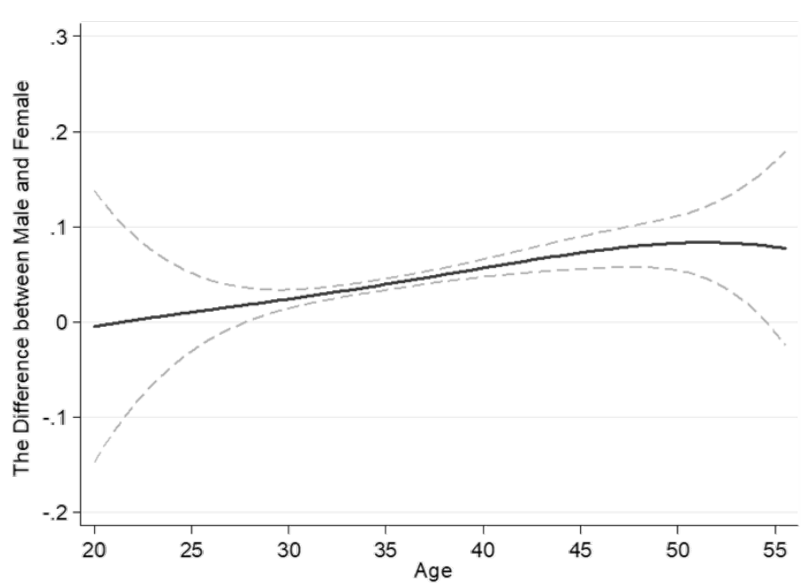


Figure B.16: Age and Marital Status Revealing as "Single"

Figure B.16a: Share of Revealing Martial Status as a Function of Age

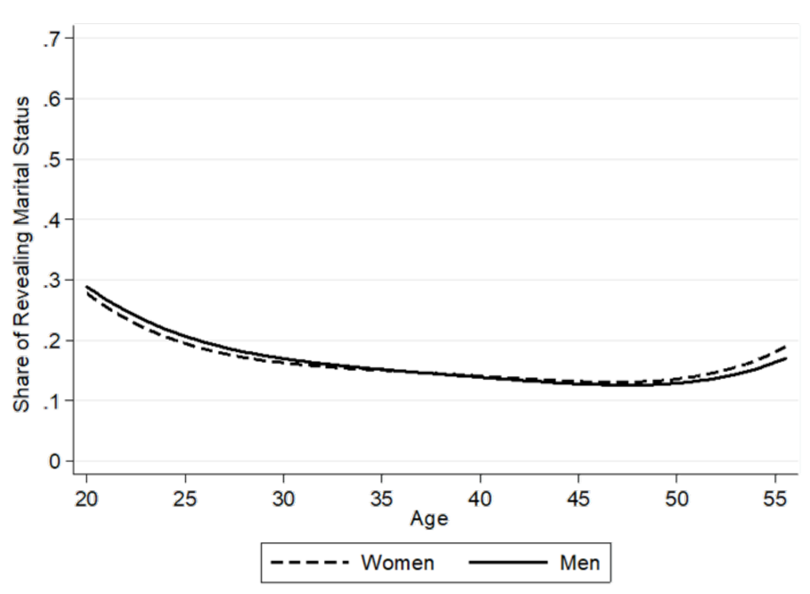



Figure B.16b: Predicted Male Share - Predicted Female Share in Revealing Martial Status

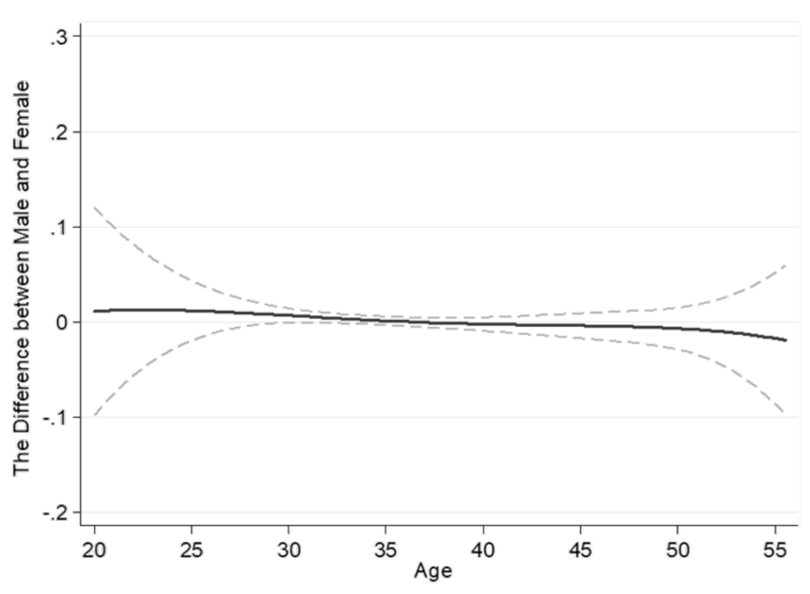


Figure B.17: Age and Marital Status Revealing as "Married"

Figure B.17a: Share of Revealing Martial Status as a Function of Age

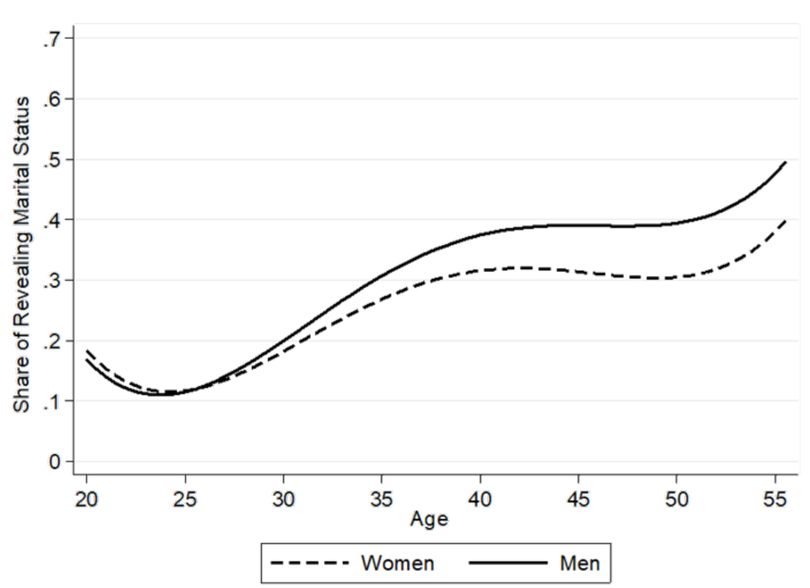
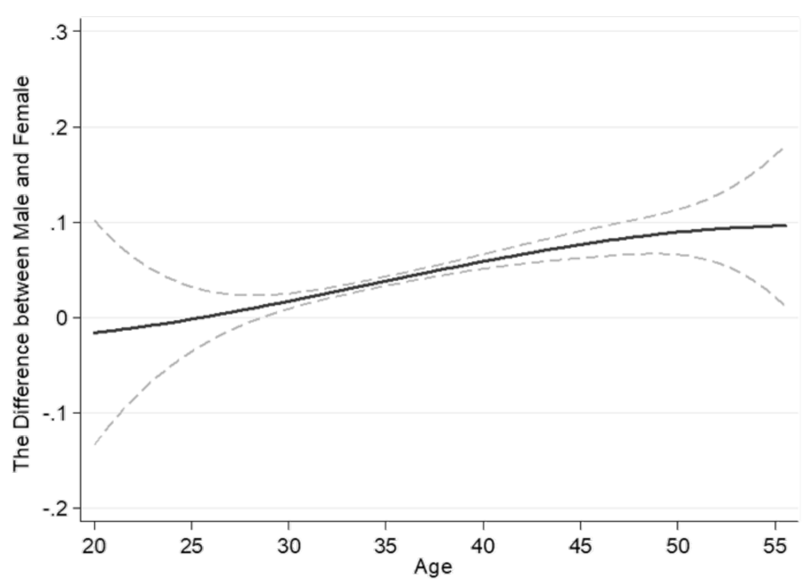



Figure B.17b: Predicted Male Share - Predicted Female Share in Revealing Martial Status



# Appendix C

# Appendix for "Measuring Algorithmic Bias in Job Recommender Systems: An Audit Study Approach"

## C.1 Resume Audit Study Experimental Design

### C.1.1 Job Type Selection

In each job board, 35 types of jobs were selected based on three criteria: the number of active job openings, the job's gender type (female-dominated jobs, gender-balanced jobs, and male-dominated jobs), and hierarchy level (entry, middle, and high). For each job type, I scraped 50 job ads to determine the education level and academic major that are required by most employers. In addition, 50 resumes in the job type were employed to derive the current wages (adjusted to be age-appropriate).

Table C.1 to C.4 list the selected job type (industry-occupation cell) in each job board, the corresponding hierarchy level (low, middle, high), the required education level, and the major. Current wage (,) represents current wages for (young, older) workers in 10k RMB, respectively.

Table C.1: Selected Job Types in Job Board 1

| Gender | Industry | Occupation | Hierarchy Level | Education Level | Major | Current Wages |
|---|---|---|---|---|---|---|
| M | Computer Software | Software Engineer | Low | Bachelor | Computer Science | (14, 17) |
| | Computer Software | Senior Software Engineer | High | Bachelor | Computer Science | (17, 23) |
| | Internet/ E-Business | Operations Specialist | Low | College | Computer Science | (7, 9) |
| | Internet/ E-Business | Operations Manager/Supervisor | High | Bachelor | Computer Science | (11, 14) |
| | Machine Manufacturing | General Worker /Operator | | College | Machinery | (8, 13) |
| | Automobiles/Motorcycles | General Worker /Operator | | College | Machinery | (9, 13) |
| | Transportation/Shipping | Courier | | College | Econ&Management | (5, 6) |
| | Internet/ E-Business | Courier | | College | Econ&Management | (6, 7) |
| | Wholesale/Retail | Warehouse Keeper | | College | Econ&Management | (4, 5) |
| N | Internet/ E-Business | Data Analyst | | Bachelor | Statistics | (11, 14) |
| | Computer Software | Data Analyst | | Bachelor | Statistics | (11, 14) |
| | Computer Software | Product Manager/Supervisor | | Bachelor | Econ&Management | (13, 17) |
| | Internet/ E-Business | Product Manager/Supervisor | | Bachelor | Econ&Management | (13, 17) |
| | Internet/ E-Business | Sales Representative | Low | College | Marketing | (5, 7) |
| | Education/Training | Sales Representative | Low | College | Marketing | (5, 7) |
| | Real Estate Services | Sales Representative | Low | College | Marketing | (6, 8) |
| | Internet/ E-Business | Sales Manager | Middle | College | Marketing | (12, 17) |
| | Computer Software | Sales Manager | Middle | College | Marketing | (12, 17) |
| | Wholesale/Retail | Sales Director | High | Bachelor | Marketing | (16, 21) |
| | Internet/ E-Business | Sales Director | High | Bachelor | Marketing | (16, 21) |
| | Internet/ E-Business | Front Desk | Low | College | Econ&Management | (6, 8) |

| | | | | | | |
|---|---|---|---|---|---|---|
| F | Professional Services | Front Desk | Low | College | Econ&Management | (6, 8) |
| | Professional Services | Executive Assistant | Low | College | Econ&Management | (7, 9) |
| | Computer Software | Executive Assistant | Low | College | Econ&Management | (7, 9) |
| | Internet/ E-Business | Executive Manager | High | College | Econ&Management | (11, 13) |
| | Wholesale/Retail | Store Clerk | Low | College | Marketing | (5, 7) |
| | Wholesale/Retail | Store Manager | High | College | Marketing | (9, 11) |
| | Internet/ E-Business | Customer Service | Low | College | Marketing | (5, 6) |
| | Finance/Securities | Customer Service | Low | College | Marketing | (5, 6) |
| | Internet/ E-Business | Customer Service Manager | High | College | Marketing | (8, 12) |
| | Trade/Import-Export | Accountant | | Bachelor | Accounting | (8, 12) |
| | Wholesale/Retail | Accountant | | Bachelor | Accounting | (8, 12) |
| | Internet/ E-Business | HR Specialist/Assistant | Low | College | Econ&Management | (6, 8) |
| | Professional Services | HR Specialist/Assistant | Low | College | Econ&Management | (6, 8) |
| | Internet/ E-Business | Human Resources Manager | High | College | Econ&Management | (9, 12) |

Table C.2: Selected Job Types in Job Board 2

| Gender | Industry | Occupation | Skill Level | Education Level | Major | Current Wages |
|---|---|---|---|---|---|---|
| M | Computer Software | Software Engineer | | Bachelor | Computer Science | (15, 23) |
| | Internet | Mobile Development Engineer | | Bachelor | Computer Science | (16, 23) |
| | Internet | Algorithm Engineer | | Bachelor | Computer Science | (17, 24) |
| | Internet | Operations Specialist | Low | College | Computer Science | (7, 9) |
| | Internet | Operations Manager/Supervisor | High | Bachelor | Computer Science | (11, 14) |
| | Real Estate Development | Real Estate Project Management | | Bachelor | Architecture | (14, 22) |
| N | Computer Software | Product Manager/Supervisor | | Bachelor | Econ&Management | (14, 20) |
| | Internet | Product Manager/Supervisor | | Bachelor | Econ&Management | (14, 20) |
| | Computer Software | Project Manager/Supervisor | | Bachelor | Econ&Management | (13, 19) |
| | Internet | Project Manager/Supervisor | | Bachelor | Econ&Management | (13, 19) |
| | Internet | Data Analyst | | Bachelor | Statistics | (12, 18) |
| | Big Data | Data Analyst | | Bachelor | Statistics | (12, 18) |
| | Securities/Investment | Data Analyst | | Bachelor | Statistics | (12, 18) |
| | Advertising/Public Relations | Public Relations Specialist/Assistant | | College | Marketing | (11, 14) |
| | Advertising/Public Relations | Public Relations Manager/Supervisor | | Bachelor | Marketing | (15, 20) |
| | E-Business | Sales Representative | Low | College | Marketing | (7, 12) |
| | Internet | Sales Representative | Low | College | Marketing | (7, 12) |
| | Education/Training | Sales Representative | Low | College | Marketing | (7, 12) |
| | Real Estate Services | Sales Representative | Low | College | Marketing | (8, 13) |
| | Wholesale/Retail | Sales Manager | Middle | College | Marketing | (12, 17) |
| | Real Estate Services | Sales Manager | Middle | College | Marketing | (12, 17) |

| | | | | | | |
|---|---|---|---|---|---|---|
| | Internet | Sales Director | High | Bachelor | Marketing | (14, 19) |
| | Wholesale/Retail | Sales Director | High | Bachelor | Marketing | (14, 19) |
| F | E-Business | Web Customer Service | Low | College | Marketing | (6, 8) |
| | Banking | Telephone Customer Service | Low | College | Marketing | (6, 8) |
| | E-Business | Customer Service Manager | High | College | Marketing | (12, 15) |
| | Banking | Customer Service Manager | High | College | Marketing | (12, 15) |
| | E-Business | Accountant | | Bachelor | Accounting | (9, 14) |
| | Internet | HR Specialist/Assistant | Low | College | Econ&Management | (6, 9) |
| | Professional Services | HR Specialist/Assistant | Low | College | Econ&Management | (6, 9) |
| | Internet | Human Resources Manager/Supervisor | High | Bachelor | Econ&Management | (11, 14) |
| | Computer Software | Human Resources Manager/Supervisor | High | Bachelor | Econ&Management | (11, 14) |
| | Internet | Executive Assistant/Secretary | Low | College | Econ&Management | (7, 9) |
| | Internet | Administration Specialist/Assistant | Low | College | Econ&Management | (6, 8) |
| | Internet | Administration Manager/Supervisor | High | College | Econ&Management | (9, 14) |

Table C.3: Selected Job Types in Job Board 3

| Gender | Industry | Occupation | Skill Level | Education Level | Major | Current Wages |
|---|---|---|---|---|---|---|
| M | Internet/E-Business | WEB Front-end Developer | | Bachelor | Computer Science | (17, 24) |
| | Machine Manufacturing | Mechanical Engineer | | Bachelor | Machinery | (16, 21) |
| | Computer Software | Software Engineer | Low | Bachelor | Computer Science | (18, 25) |
| | Computer Software | Senior Software Engineer | High | Bachelor | Computer Science | (22, 27) |
| | Internet/E-Business | Operations Specialist | Low | College | Computer Science | (10, 13) |
| | Internet/E-Business | Operations Manager/Supervisor | High | Bachelor | Computer Science | (14, 20) |
| | Real Estate Development | Architect | | Bachelor | Architecture | (15, 22) |
| N | Pharmaceuticals/Biotechnology | Sales Representative | Low | College | Marketing | (10, 15) |
| | Securities/Investment Funds | Sales Representative | Low | College | Marketing | (11, 15) |
| | Pharmaceuticals/Biotechnology | Sales Manager/Supervisor | Middle | Bachelor | Marketing | (14, 18) |
| | Internet/E-Business | Sales Manager/Supervisor | Middle | Bachelor | Marketing | (13, 18) |
| | Securities/Investment Funds | Sales Manager/Supervisor | Middle | Bachelor | Marketing | (13, 18) |
| | Pharmaceuticals/Biotechnology | Sales Director | High | Bachelor | Marketing | (17, 24) |
| | Internet/E-Business | Sales Director | High | Bachelor | Marketing | (16, 25) |
| | Commodity | Sales Director | High | Bachelor | Marketing | (16, 24) |
| | Internet/E-Business | Product Manager/Supervisor | | Bachelor | Econ&Management | (15, 22) |
| | Computer Software | Product Manager/Supervisor | | Bachelor | Econ&Management | (15, 22) |
| | Internet/E-Business | Project Manager/Supervisor | | Bachelor | Econ&Management | (15, 22) |
| | Computer Software | Project Manager/Supervisor | | Bachelor | Econ&Management | (15, 22) |
| | Commodity | Marketing Manager/Supervisor | | Bachelor | Marketing | (14, 22) |
| | Wholesale/Retail | Marketing Manager/Supervisor | | Bachelor | Marketing | (14, 22) |

| | | | | | | |
|---|---|---|---|---|---|---|
| | Real Estate Development | Legal manager/Supervisor | | Bachelor | Law | (15, 25) |
| | Internet/E-Business | Legal manager/Supervisor | | Bachelor | Law | (15, 24) |
| F | Internet/E-Business | Human Resources Specialist/Assistant | Low | College | Econ&Management | (9, 12) |
| | Real Estate Development | Human Resources Specialist/Assistant | Low | College | Econ&Management | (9, 12) |
| | Internet/E-Business | Human Resources Manager/Supervisor | Middle | Bachelor | Econ&Management | (14, 20) |
| | Real Estate Development | Human Resources Manager/Supervisor | Middle | Bachelor | Econ&Management | (14, 20) |
| | Internet/E-Business | Human Resources Director | High | Bachelor | Econ&Management | (16, 26) |
| | Real Estate Development | Human Resources Director | High | Bachelor | Econ&Management | (16, 26) |
| | Internet/E-Business | Accountant | Low | Bachelor | Accounting | (12, 18) |
| | Securities/Investment Funds | Financial Manager | High | Bachelor | Finance | (15, 20) |
| | Internet/E-Business | Administration Specialist/Assistant | Low | College | Econ&Management | (9, 13) |
| | Real Estate Development | Executive Assistant/Secretary | Low | College | Econ&Management | (10, 14) |
| | Internet/E-Business | Administration Manager/Supervisor | Low | Bachelor | Econ&Management | (15, 20) |
| | Internet/E-Business | Administration Vice President | High | Bachelor | Econ&Management | (51, 88) |

Table C.4: Selected Job Types in Job Board 4

| Gender | Occupation | Skill Level | Education Level | Major | Current Wages |
|---|---|---|---|---|---|
| M | WEB Front-end Developer | | Bachelor | Computer Science | (19, 25) |
| | Operation and Maintenance Engineer | Low | Bachelor | Computer Science | (18, 24) |
| | Operation and Maintenance Director | High | Bachelor | Computer Science | (19, 26) |
| | Pattern Recognition | | Bachelor | Computer Science | (19, 25) |
| | Machine Learning | | Bachelor | Computer Science | (19, 25) |
| | Operations Assistant | Low | College | Computer Science | (7, 9) |
| | Operations Specialist | Middle | College | Computer Science | (10, 12) |
| | Operations Manager/Supervisor | High | Bachelor | Computer Science | (14, 19) |
| | Test Engineer | Low | Bachelor | Computer Science | (15, 22) |
| | Test Manager | High | Bachelor | Computer Science | (19, 25) |
| | Data Architect | | Bachelor | Computer Science | (17, 25) |
| N | Sales Representative | Low | College | Marketing | (8, 12) |
| | Sales Manager/Supervisor | Middle | Bachelor | Marketing | (13, 17) |
| | Sales Director | High | Bachelor | Marketing | (18, 25) |
| | Product Assistant | Low | College | Econ&Management | (9, 10) |
| | Product Manager | High | Bachelor | Econ&Management | (15, 23) |
| | Project Assistant | Low | College | Econ&Management | (9, 10) |
| | Project Manager | High | Bachelor | Econ&Management | (15, 23) |
| | Data Analyst | | Bachelor | Statistics | (13, 19) |
| | Design Assistant | Low | College | Arts | (8, 10) |
| | Designer | Middle | College | Arts | (13, 19) |

| | | | | | |
|---|---|---|---|---|---|
| | Design Manager | High | Bachelor | Arts | (15, 23) |
| | Strategy Consultant | | Bachelor | Econ&Management | (13, 19) |
| F | Human Resources Specialist/Assistant | Low | College | Econ&Management | (9, 10) |
| | Human Resources Manager/Supervisor | Middle | Bachelor | Econ&Management | (14, 20) |
| | Human Resources Director | High | Bachelor | Econ&Management | (17, 26) |
| | Accountant | Low | Bachelor | Accounting | (13, 17) |
| | Training Specialist | | College | Econ&Management | (10, 12) |
| | Customer Service | Low | College | Marketing | (7, 8) |
| | Customer Service Manager | High | College | Marketing | (13, 17) |
| | Media Specialist | Low | College | Marketing | (7, 8) |
| | Media Manager | High | Bachelor | Marketing | (10, 15) |
| | Administration Specialist/Assistant | Low | College | Econ&Management | (9, 12) |
| | Administration Manager/Supervisor | Middle | Bachelor | Econ&Management | (13, 18) |
| | Administration Director | High | Bachelor | Econ&Management | (16, 25) |

Notes: The industry in job board 4 is set as "all industries".

### C.1.2 Fictitious Resume

The resumes only contain the basic information required by each job board to register as a valid job seeker. The first section of a fictitious resume is personal information, including worker's name, birth date, years of working experience, current wage, city, employment status, phone number, and email address. The second part is about worker's education: the highest education level, time period, university name and major. The third part describes worker's working experience of the most recent job including the time period, company name, occupation, industry, job title, and job description. The last part is worker's intention for future jobs, including desired wage, desired location, desired industry, and occupation. Two workers in each gender pair have identical backgrounds, and four workers in each group (young male, young female, older male, older female) are placed in each job type.

#### Personal Information

*Name*: I picked up the most popular first and last names to make up the names of fictitious applicants. Based on the statistics from 2015 Chinese Census 1% Population Sample, I chose the top 20 last names, top 15 male first names, and top 15 female first names as the applicants' name pool (listed in Appendix C.1.2). For each applicant, the last name and first name corresponding to the applicant's gender will be randomly drawn from the name pool. Although gender is explicitly stated in the resume and we do not need applicant's name to denote gender, I still adopted first names that are consistent with a worker's gender to make the fictitious profile as common and real as possible.

**Names of Fictitious Applicants**

*Last name*: 李(Li), 王(Wang), 张(Zhang), 刘(Liu), 陈(Chen), 杨(Yang), 赵(Zhao),

黄(Hunag), 周(zhou), 吴(Wu), 徐(Xu), 孙(Sun), 胡(Hu), 朱(Zhu), 高(Gao), 林(Lin), 何(He), 郭(Guo), 马(Ma), 罗(Luo).

*Male First Name*: 伟(Wei), 强(Qiang), 磊(Lei), 军(Jun), 洋(Yang), 勇(Yong), 杰(Jie), 涛(Tao), 超(Chao), 平(Ping), 刚(Gang), 浩(Hao), 鹏(Peng), 宇(Yu), 明(Ming).

*Female First Name*: 芳(Fang), 娜(Na), 敏(Min), 静(Jing), 丽(Li), 艳(Yan), 娟(Juan), 霞(Xia), 婷(Ting), 雪(Xue), 丹(Dan), 英(Ying), 洁(Jie), 玲(Ling), 燕(Yan).

*Birth Date*: Employers infer worker's age from the birth date. Instead of varying workers' age directly, I used their graduation year to classify the age level, and "older" workers refer to ones who graduated earlier and have more working experience. Applicants have two potential age levels: Young workers graduated in 2017, and old workers graduated in 2007. After a worker's graduation year is fixed, his age is jointly determined by the graduation year and his education level. The advantage of this design is that workers' years of working experience are equalized within each age level. More specifically, young workers are 25 (with a college degree, born in 1995) or 26 (with a bachelor's degree, born in 1994) with three years of working experience, 35 or 36 years old are for the senior workers with more than 5 years of working experience. Workers in the gender pair have the same randomly drawn birth month and day.

*Years of Working Experience*: To simplify the profiles, I assumed workers started to work just after they graduated from the university/college of their highest degree. As discussed above, years of working experience is the difference between the current year (2020) and the graduation year. For instance, if a worker graduated in 2017, then he has 2020 – 2017, three years of working experience.

*Current Wage*: Fictitious workers' wages are drafted based on wages of active workers in job boards by matching their current job position as well as working experience. I used the hiring agent account in each platform and searched for workers that were currently in the job positions and specified the working experience as "1 to 3 years" and "5 to 10 years" in March 2020. For each experience level in every job position, I recorded the first 50 workers' current wages shown in the search result and took the average as the fictitious worker's wage.

*City*: All of the four job boards are nationally recognized and cover most of the regions in China, and over half of job postings are from first-tier cities. To achieve enough amount of job recommendations, fictitious workers are currently living in the first-tier cities, including Beijing, Shanghai, Shenzhen, and Guangzhou. Employment Status: All of the workers are currently employed.

*Phone number and email*: Each applicant has a unique and active email address and mobile phone number.

**Education**

Workers' education level is designed to match jobs' education requirements. For each job type, I checked 100 job advertisements in February 2020 and listed the most common education. 85% of job ads required workers had a bachelor's or junior college degree. Bachelor's degree often takes 4 years to achieve, while junior college takes 3 years. The end time of school is the graduation year, and the start time of school depends on worker's education degree, which is three years (college degree) or four years (bachelor's degree) earlier than the graduation year. For instance, a young worker, graduated with a bachelor's degree in June 2017, is 26 years old (born in 1994) and started his university program in August 2013.

Two workers in the same gender pair have the same educational background, and the school's name is randomly drawn from the Chinese High Education Institution List, released by the Ministry of Education in 2019. Majors will also match job positions: Computer Science/Software is for IT jobs, Mathematics/Statistics is for data position, and economics/management/marketing majors are for other jobs.

### Recent Job History

As we assume all the workers are currently employed, their recent jobs are their current jobs. For young workers, their current jobs started in August in the year when they graduated with the highest degree (2017); for old workers, their current jobs started five years ago, in March 2015, implying that they have 5 years tenure in their recent positions.

I made up company names to minimize the disturbance to both job seekers and employers on job platforms. The company name consists of three parts: (1) company's location. It will be the same with worker's current city. (2) company's name. I used an online business name generator to collect 100 company names listed below. The company name will be randomly assigned to each gender pair. (3) company's industry. It will be consistent with the job's industry. An example of the company name is, Beijing Dongya Internet Technology Company. Worker's current occupation and industry will be the same as the job's occupation and industry. Job title and job description are filled in by words, and I set them as the job's occupation.

**Names of Company**

东艾, 森利, 先卓, 利晟, 同通, 富长盛, 芯达, 精典, 尼佳, 益复捷, 生德, 晶长, 森益, 金伙伴, 德光, 茂全, 鲜派, 信顺康, 龙丝, 新耀协, 佳丽, 晖, 佳洲, 森道尔, 皇祥千, 润飞昌, 福中荣, 基玉, 如和, 茂乾, 翔鹏, 南湘, 圣泰, 吉春, 本寿, 亚义金, 耀浩, 邦洁, 宝复, 洪进贵, 永泰满, 显郦, 华行, 韵仪, 格派, 晶佩, 迪和, 领速, 贝耀, 信华诚, 世力, 舜杰, 久福, 曼新, 仁大兴, 金祥元, 泰伟飞, 亚和金, 吉振, 和伟中, 盛金缘, 立韦, 宏久, 吉至, 曼展, 天联, 金涛, 网诚, 系广, 圣金龙, 易露发, 嘉利华, 聚顿, 公同宏, 威邦, 力涛, 恒蓝, 铭航, 中美公, 永逸, 同捷, 发和, 易龙, 汉金, 干亚, 翔洋, 新都, 茂进永, 达通, 娇罗, 浩中和, 东升, 龙姿, 隆新弘, 仟顺, 越福, 川实, 中协吉, 霸辉, 洪谦, 裕飞

**Job Intention**

A worker searches for full-time jobs, in which the desired wage is 120% of his current wage (or the wage range), and the desired city, industry, and occupation will be the same as the current ones. The table below summarizes the information included in worker's resume.

Table C.5: Resume Information Generation

| | Method | Note |
|---|---|---|
| **Personal Information** | | |
| Name | Randomly assigned to each worker | |
| Birth Date | Young worker graduated in 2017, and older worker graduated in 2007. Birth year is decided by graduation year and education level. | Young, bachelor's =1994,<br>Young, college=1995.<br>Older, bachelor's =1984,<br>Older, college=1985. |
| Years of Working Experience | 2020 - graduation year | 3 or 13 years |
| Current Wage | Average wage of the collected workers in the platforms. | Adjust with job type and experience. |
| City | Beijing, Shanghai, Shenzhen, Guangzhou | |
| Employment Status | Currently employed. | |
| Phone Number & Email | Uniquely assigned for each worker. | |
| **Education** | | |
| Highest degree | Assigned on group level, based on job type's education requirement. | Bachelor's degree or junior college. |
| Time Period | Graduation year – years to achieve the highest degree. | 4 years to achieve bachelor's degree,<br>3 years to achieve college degree. |
| School Name | Randomly drawn for each gender pair. | Chinese High Education Institution List (2019) |
| Major | Same on group level. | Depends on job type. |
| **Recent Job** | | |
| Time Period | Young worker: after graduation (2017) until now, | |

| | | |
|---|---|---|
| | Older worker: 2015 until now. | |
| Company Name | Location +name + industry, name will be randomly assigned to each worker. | |
| Occupation | Same with job type | |
| Industry | Same with job type | |
| Job Title | Same with occupation | |
| Job Description | Same with occupation | |
| **Intention** | | |
| Desired Wage | Current wage*1.2 | |
| Desired City | Same with city | |
| Desired Industry | Same with job type | |
| Desired Occupation | Same with job type | |

## C.2 Robustness Check on Set Difference of Job Recommendations

Table C.6: Gender Differences on Explicit Measures of Recommended Jobs

| | Male – Female | |
|---|---|---|
| | Two-sample t-test with Unequal Variance | Wilcoxon Rank-sum Test |
| Posted Wage | 2708.78* | 2.0474** |
| | (1527.46) | |
| Required Education | 0.0226 | 0.8415 |
| | (0.0259) | |
| Required Experience | 0.0779*** | 3.5723*** |
| | (0.0262) | |

Notes:

1. Gender difference is computed from the mean of male-only jobs minus the mean of female-only jobs.
2. In column 2, standard errors are in parentheses, which are derived from two-sample t-tests with unequal variance.
3. Column 3 reports z-value from Wilcoxon Rank-sum Test.
4. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

## C.3 List Difference of Job Recommendations

Table C.7: List Difference Rate in Job Recommendations

| | Share |
|---|---|
| All | 0.7069 |
| Round | |
| 0 | 0.5827 |
| 1 | 0.8642 |
| 2 | 0.8634 |
| 3 | 0.8643 |
| Age | |
| Young | 0.7048 |
| Old | 0.7092 |
| Gender | |
| Female | 0.7039 |
| Neutral | 0.7093 |
| Male | 0.7079 |
| Hierarchy | |
| Entry | 0.7049 |
| Middle | 0.7077 |
| High | 0.7083 |
| City | |
| Beijing | 0.7018 |
| Shanghai | 0.7090 |
| Shenzhen | 0.7104 |
| Guangzhou | 0.7067 |

Notes:

1. List difference is defined as the share of different recommendations in the recommendation list, as shown in Figure 2b.

Table C.8: Gender (List) Differences on Explicit Measures of Recommended Jobs

| | Male – Female | |
|---|---|---|
| | Paired t-test | Wilcoxon Signed Rank Test |
| Posted Wage | 1207.55 | 1.3348 |
| | (830.11) | |
| Required Education | 0.0084 | 0.7739 |
| | (0.0093) | |
| Required Experience | 0.0042 | 0.5489 |
| | (0.0090) | |

Notes:

1. Gender difference is computed from the mean of male-only jobs minus the mean of female-only jobs.
2. In column 2, standard errors are in parentheses, which are derived from paired-sample t-tests.
3. Column 3 reports z-value from Wilcoxon Signed Rank Test.
4. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

## C.4 Words in Job Recommendations

Figure C.1: Word Cloud from Job Ads

Table C.9: Resume Information Generation

| | |
|---|---|
| Skills | Literacy skills: listen, speak, read, write, language, documentation<br>Numeracy skills: data, accounting, analysis<br>ICT skills: programing, microsoft office, chat tools<br>Problem solving skills: learning, comprehension, thinking, logic, decision-making, planning, problem-solving, engineering, independent<br>Influencing skills: leadership, team management, charge, supervise<br>Co-operative skills: cooperation, communication, teamwork, assist, coordination, organize, negotiate, public relation, marketing, sale, client, compliance<br>Self-organizing skills: administrative, design, collect, reception, driving, execution, test, task management |
| Work Form | Schedule: work shift, night work, morning work, evening work, big and small week*, eight-hour, flexible, attendance, overtime, no overtime<br>Business travel: regular travel, short travel, long travel<br>Work break: weekly break, monthly break, noon break, regular working hour |
| Benefits | Payment: base pay, commission, stock, allowance, promotion, reward<br>Break: vacation, marriage leave, parental leave, maternity leave, sick leave, funeral leave, holiday<br>Facilities: office supplements, vehicle, meal, housing, shuttle, nearby, commute<br>Insurance: fiveone*, medical insurance, commercial insurance, social security, funds, maternity insurance, unemployment insurance, endowment insurance, injury insurance, disease insurance<br>Other benefits: training, staffing, activities, mentor |
| Company | Environment: atmosphere, employee care, career, dream, culture, screening<br>Type: direct recruiting, public company, top500, startup, flat, financing, big company* |

| | |
|---|---|
| | Title: senior, medium, core |
| Requirements | Education: non education, certificate, new grad, tongzhao*, tier-one school, fulltime school, top school, nonmajor, major, science&engineering<br>Experience: non experience, experienced, oversea<br>Demographics: non gender, non age, below35, below40<br>Personality: effective, rigorous, carefully, patient, energetic, active, outgoing, optimistic, virtuous, trustworthy, honest, practical, self-motivated, hardworking, passion, tenacious, sharp mind, generous, curious, courageous, innovative, punctual, entrepreneurial, devotion, enthusiasm, kind, responsibility, pressure<br>Appearance: figure, temperament, healthy, facial, clothing, shape<br>Objective: voice, responsive, no crime, regulation, solitary |

Notes:

1. Table C.9 shows the extracted words from job ads in four job boards, and the restrictions are described in Section 6.2.2.
2. Every listed word includes its variations on parts of speech, such as leadership vs leading, and confidence vs confident.
3. "fiveone" represents "five social insurance and one housing fund" (五险一金), including endowment insurance, medical insurance, unemployment insurance, employment injury insurance, maternity insurance, and housing fund. Big and small week describes the working schedule in which workers have one-day rest in one week and two-day rest in the next week. Big company indicates companies that have more than 1000 employees. Tongzhao means university or college admission is through Gaokao in high school.

### C.4.1 Survey from Amazon MTurk

To determine the gendered perceptions of words, I recruited participants from Amazon's Mechanical Turk (MTurk) in September 2021 to choose whether the existence of a certain word in the job ad indicates gender stereotypes and implicit gender preferences of employers.

The survey question is: "Suppose you are the hiring agent of a company, and plan to post a job advertisement that contains the word X in the job description. This indicates that you prefer to hire (1) no gender request for worker; (2) male worker; (3) female worker". In total, 86 valid surveys were collected from people between the ages of 25 to 55, and 56% of them were men. The gender score of a word is computed as:

*Score = -1*number of participants choose* (3) + *1* number of participants choose* (2)

, in which -1 indicates the extreme female word and 1 implies the extreme male word. The average gender score of words in the survey is 0.0905 and the standard deviation is 0.1111. Male words are defined as words whose scores are above one standard deviation from the mean, 0.2016, and female words' scores are below one standard deviation from the mean, -0.0206.[1]

[1] tierone university and tongzhao are excluded from the surveyed words because they are only identified in the Chinese high-level education system.

Table C.10: Gendered Words from Amazon MTurk Survey

| Female Words | Male Words |
|---|---|
| administrative, assist, careful, compliance, design, documentation, enthusiasm, figure, holiday, kind, learning, marriage leave, maternity insurance, maternity leave, parental leave, patient, read, reception, shape, sick leave, temperament, voice, writing | analysis, big week, commission, data, driving, responsibility, effective, engineering, evening work, experienced, independent, no crime, leadership, logic, long travel, mentor, negotiation, nightwork, overtime, practical, pressure, promotion, science&engineering, startup, stock, supervise, vehicle, work shift |

### C.4.2 Survey from Amazon MTurk

The Chinese version of the survey on people's perceptions about gendered words in job ads was conducted in Wenjuanxing (问卷星) in September 2021. The surveyed question is the same as the one from AMturk, but in Chinese: 假设您是公司HR，发布的招聘广告中包含以下词汇，代表您倾向于招聘(1) 性别不限; (2) 男员工; (3) 女员工。

79 valid respondents participated in the survey, 81% of them were between 25 to 55 years old and 73% of them were men. The average gender score of words in the survey is 0.0962 and the standard deviation is 0.0721. Male words are defined as words whose scores are above one standard deviation from the mean, 0.1683, and female words' scores are below one standard deviation from the mean, 0.0241.

Table C.11: Gendered Words from Chinese Survey

| Female Words | Male Words |
|---|---|
| active, administrative, assist, atmosphere, care, collect, communication, compliance, design, eight hour, facial, figure, flexible, health, kind, listen, marriage leave, maternity ins, office supplements, outgoing, parental leave, passion, patient, read, reception, shape, sick leave, speak, temperament, voice, writing | below40, charge, commission, commute, core, courageous, culture, data, disease ins, driving, responsibility, engineering, enterprise, independent, injury ins, no crime, leadership, long travel, meal, negotiation, nightwork, optimistic, oversea, overtime, practical, pressure, promotion, responsive, screening, self-motivated, solving, staffing, stock, teamwork, tenacious, training, unemployment ins, punctual |

# Bibliography


Agan, A., Cowgill, B. and Gee, L. (2021), Salary history and employer demand: Evidence from a two-sided audit, Technical report, National Bureau of Economic Research.

Agan, A., Cowgill, B. and Gee, L. K. (2020), Do workers comply with salary history bans? a survey on voluntary disclosure, adverse selection, and unraveling, *in* 'AEA Papers and Proceedings', Vol. 110, pp. 215–19.

Agan, A. and Starr, S. (2017), 'The effect of criminal records on access to employment', *American Economic Review* **107**(5), 560–64.

Ahmed, T. (2018), 'Employment growth/skill requirement estimation in india: a non-traditional approach', *AUSTRALIAN JOURNAL OF LABOUR ECONOMICS* **21**(3).

Al-Otaibi, S. T. and Ykhlef, M. (2012), 'A survey of job recommender systems', *International Journal of Physical Sciences* **7**(29), 5127–5142.

Albert, R., Escot, L. and Fernández-Cornejo, J. A. (2011), 'A field experiment to study sex and age discrimination in the madrid labour market', *The International Journal of Human Resource Management* **22**(02), 351–375.

Albrecht, J., Gautier, P. A. and Vroman, S. (2006), 'Equilibrium directed search with multiple applications', *The Review of Economic Studies* **73**(4), 869–891.

Ali, M., Sapiezynski, P., Bogen, M., Korolova, A., Mislove, A. and Rieke, A. (2019), 'Discrimination through optimization: How facebook's ad delivery can lead to biased outcomes', *Proceedings of the ACM on Human-Computer Interaction* **3**(CSCW), 1–30.

Angwin, J., Larson, J., Mattu, S. and Kirchner, L. (2016), 'Machine bias', *ProPublica, May* **23**(2016), 139–159.

Angwin, J., Scheiber, N. and Tobin, A. (2017), 'Facebook job ads raise concerns about age discrimination', *The New York Times* **20**, 1.

Arnold, D., Dobbie, W. and Hull, P. (2021), Measuring racial discrimination in algorithms, *in* 'AEA Papers and Proceedings', Vol. 111, pp. 49–54.

Athreya, B. (2021), 'Bias in, bias out: gender and work in the platform economy'.

Avivi, H., Kline, P., Rose, E. and Walters, C. (2021), Adaptive correspondence experiments, *in* 'AEA Papers and Proceedings', Vol. 111, pp. 43–48.

Baert, S. (2018), 'Hiring discrimination: An overview of (almost) all correspondence experiments since 2005', *Audit studies: Behind the scenes with theory, method, and nuance* pp. 63–77.

Baert, S., Albanese, A., du Gardein, S., Ovaere, J. and Stappers, J. (2017), 'Does work experience mitigate discrimination?', *Economics Letters* **155**, 35–38.

Baker, M. and Cornelson, K. (2018), 'Gender-based occupational segregation and sex differences in sensory, motor, and spatial aptitudes', *Demography* **55**(5), 1749–1775.

Baker, M., Halberstam, Y., Kroft, K., Mas, A. and Messacar, D. (2019), Pay transparency and the gender gap, Technical report, National Bureau of Economic Research.

Ballance, J., Clifford, R. and Shoag, D. (2020), '"no more credit score": Employer credit check bans and signal substitution', *Labour Economics* **63**, 101769.

Banfi, S., Choi, S. and Villena-Roldan, B. (2019), 'Deconstructing job search behavior', *Available at SSRN 3323545* .

Banfi, S. and Villena-Roldan, B. (2019), 'Do high-wage jobs attract more applicants? directed search evidence from the online labor market', *Journal of Labor Economics* **37**(3), 715–746.

Barach, M. A. and Horton, J. J. (2021), 'How do employers use compensation history? evidence from a field experiment', *Journal of Labor Economics* **39**(1), 193–218.

Barocas, S. and Selbst, A. D. (2016), 'Big data's disparate impact', *Calif. L. Rev.* **104**, 671.

Bartik, A. W. and Nelson, S. T. (2016), *Credit reports as resumes: The incidence of pre-employment credit screening*, Massachusetts Institute of Technology, Department of Economics.

Bartlett, R., Morse, A., Stanton, R. and Wallace, N. (2021), 'Consumer-lending discrimination in the fintech era', *Journal of Financial Economics* .

Barzilay, A. R. and Ben-David, A. (2016), 'Platform inequality: gender in the gig-economy', *Seton Hall L. Rev.* **47**, 393.

Bedard, K. and Kuhn, P. (2015), 'Micro-marketing healthier choices: Effects of personalized ordering suggestions on restaurant purchases', *Journal of health economics* **39**, 106–122.

Bederson, B. B., Jin, G. Z., Leslie, P., Quinn, A. J. and Zou, B. (2018), 'Incomplete disclosure: Evidence of signaling and countersignaling', *American Economic Journal: Microeconomics* **10**(1), 41–66.

Behaghel, L., Crépon, B. and Le Barbanchon, T. (2015), 'Unintended effects of anonymous resumes', *American Economic Journal: Applied Economics* **7**(3), 1–27.

Belot, M., Kircher, P. and Muller, P. (2017), 'How wage announcements affect job search behaviour-a field experimental investigation', *European University Institute, working paper* .

Bem, S. L. (1981), 'Gender schema theory: A cognitive account of sex typing.', *Psychological review* **88**(4), 354.

Bennedsen, M., Simintzi, E., Tsoutsoura, M. and Wolfenzon, D. (2019), Do firms respond to gender pay gap transparency?, Technical report, National Bureau of Economic Research.

Bertrand, M. and Duflo, E. (2017), 'Field experiments on discrimination', *Handbook of economic field experiments* **1**, 309–393.

Bertrand, M., Goldin, C. and Katz, L. F. (2010), 'Dynamics of the gender gap for young professionals in the financial and corporate sectors', *American economic journal: applied economics* **2**(3), 228–55.

Bertrand, M. and Mullainathan, S. (2004), 'Are emily and greg more employable than lakisha and jamal? a field experiment on labor market discrimination', *American economic review* **94**(4), 991–1013.

Beutel, A., Chen, J., Doshi, T., Qian, H., Wei, L., Wu, Y., Heldt, L., Zhao, Z., Hong, L., Chi, E. H. et al. (2019), Fairness in recommendation ranking through pairwise comparisons, *in* 'Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining', pp. 2212–2220.

Black, S. E. and Spitz-Oener, A. (2010), 'Explaining women's success: technological change and the skill content of women's work', *The Review of Economics and Statistics* **92**(1), 187–194.

Blau, D. M. and Robins, P. K. (1990), 'Job search outcomes for the employed and unemployed', *Journal of political Economy* **98**(3), 637–655.

Board, O. (2009), 'Competition and disclosure', *The Journal of Industrial Economics* **57**(1), 197–213.

Booth, A. L. (2009), 'Gender and competition', *Labour Economics* **16**(6), 599–606.

Booth, A. and Leigh, A. (2010), 'Do employers discriminate by gender? a field experiment in female-dominated occupations', *Economics Letters* **107**(2), 236–238.

Bos, M., Breza, E. and Liberman, A. (2018), 'The labor market effects of credit market information', *The Review of Financial Studies* **31**(6), 2005–2037.

Bozdag, E. (2013), 'Bias in algorithmic filtering and personalization', *Ethics and information technology* **15**(3), 209–227.

Brenčič, V. (2010), 'Do employers respond to the costs of continued search?', *Oxford Bulletin of Economics and Statistics* **72**(2), 221–245.

Brenčič, V. (2012), 'Wage posting: evidence from job ads', *Canadian Journal of Economics/Revue canadienne d'économique* **45**(4), 1529–1559.

Brenčič, V. and Norris, J. B. (2012), 'Employers'on-line recruitment and screening practices', *Economic Inquiry* **50**(1), 94–111.

Burdett, K. (1978), 'A theory of employee job search and quit rates', *The American Economic Review* **68**(1), 212–220.

Burke, R., Sonboli, N. and Ordonez-Gauger, A. (2018), Balanced neighborhoods for multi-sided fairness in recommendation, *in* 'Conference on Fairness, Accountability and Transparency', PMLR, pp. 202–214.

Bustelo, M., Díaz Escobar, A. M., Lafortune, J., Piras, C., Salas, L. M. and Tessada, J. A. (2020), What is the price of freedom? estimating women's willingness to pay for job schedule flexibility, Technical report, IDB Working Paper Series.

Cahuc, P., Postel-Vinay, F. and Robin, J.-M. (2006), 'Wage bargaining with on-the-job search: Theory and evidence', *Econometrica* **74**(2), 323–364.

California Legislature (2017), 'Assembly bill no. 168, chapter 688'. https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201720180AB168.

Caliskan, A., Bryson, J. J. and Narayanan, A. (2017), 'Semantics derived automatically from language corpora contain human-like biases', *Science* **356**(6334), 183–186.

Capéau, B., Eeman, L., Groenez, S., Lamberts, M. et al. (2012), 'Two concepts of discrimination: inequality of opportunity versus unequal treatment of equals', *Ecore Discussion Papers* **58**, 127.

Card, D., Cardoso, A. R., Heining, J. and Kline, P. (2018), 'Firms and labor market inequality: Evidence and some theory', *Journal of Labor Economics* **36**(S1), S13–S70.

Card, D., Cardoso, A. R. and Kline, P. (2016), 'Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women', *The Quarterly journal of economics* **131**(2), 633–686.

Carlsson, M. (2011), 'Does hiring discrimination cause gender segregation in the swedish labor market?', *Feminist Economics* **17**(3), 71–102.

Carlsson, M. and Rooth, D.-O. (2007), 'Evidence of ethnic discrimination in the swedish labor market using experimental data', *Labour economics* **14**(4), 716–729.

Carrington, W. J. and Troske, K. R. (1997), 'On measuring segregation in samples with small units', *Journal of Business & Economic Statistics* **15**(4), 402–409.

Cediey, E. and Foroni, F. (2008), *Discrimination in access to employment on grounds of foreign origin in France: A national survey of discrimination based on the testing methodology of the International Labour Office*, ILO.

Chalfin, A., Danieli, O., Hillis, A., Jelveh, Z., Luca, M., Ludwig, J. and Mullainathan, S. (2016), 'Productivity and selection of human capital with machine learning', *American Economic Review* **106**(5), 124–27.

Chaturvedi, S., Mahajan, K., Siddique, Z. et al. (2021), Words matter: Gender, jobs and applicant behavior, Technical report, IZA DP No. 14497.

Chen, L., Ma, R., Hannák, A. and Wilson, C. (2018), Investigating the impact of gender on rank in resume search engines, *in* 'Proceedings of the 2018 chi conference on human factors in computing systems', pp. 1–14.

Christl, M. and Köppl-Turyna, M. (2020), 'Gender wage gap and the role of skills and tasks: evidence from the austrian piaac data set', *Applied Economics* **52**(2), 113–134.

Conlin, M., Dickert-Conlin, S. and Chapman, G. (2013), 'Voluntary disclosure and the strategic behavior of colleges', *Journal of Economic Behavior & Organization* **96**, 48–64.

Cook, C., Diamond, R., Hall, J. V., List, J. A. and Oyer, P. (2021), 'The gender earnings gap in the gig economy: Evidence from over a million rideshare drivers', *The Review of Economic Studies* **88**(5), 2210–2238.

Correll, S. J., Benard, S. and Paik, I. (2007), 'Getting a job: Is there a motherhood penalty?', *American journal of sociology* **112**(5), 1297–1338.

Cortés, P., Pan, J., Pilossoph, L. and Zafar, B. (2021), Gender differences in job search and the earnings gap: Evidence from business majors, Technical report, National Bureau of Economic Research.

Cowgill, B. (2018), 'Bias and productivity in humans and algorithms: Theory and evidence from resume screening', *Columbia Business School, Columbia University* **29**.

Craswell, N., Zoeter, O., Taylor, M. and Ramsey, B. (2008), An experimental comparison of click position-bias models, *in* 'Proceedings of the 2008 international conference on web search and data mining', pp. 87–94.

Cullen, Z. B. and Perez-Truglia, R. (2020), 'The salary taboo: Privacy norms and the diffusion of information', *National Bureau of Economic Research* (No. 25145).

Cutler, D. M., Glaeser, E. L. and Vigdor, J. L. (1999), 'The rise and decline of the american ghetto', *Journal of political economy* **107**(3), 455–506.

de Sá, C. C., Gonçalves, M. A., Sousa, D. X. and Salles, T. (2016), Generalized broof-l2r: a general framework for learning to rank based on boosting and random forests, *in* 'Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval', pp. 95–104.

Depoers, F. (2000), 'A cost benefit study of voluntary disclosure: some empirical evidence from french listed companies', *European Accounting Review* **9**(2), 245–263.

DeVaro, J. and Kauhanen, A. (2016), 'An "opposing responses" test of classic versus market-based promotion tournaments', *Journal of Labor Economics* **34**(3), 747–779.

Doleac, J. L. and Hansen, B. (2020), 'The unintended consequences of "ban the box": Statistical discrimination and employment outcomes when criminal histories are hidden', *Journal of Labor Economics* **38**(2), 321–374.

Dreisbach, C., Koleck, T. A., Bourne, P. E. and Bakken, S. (2019), 'A systematic review of natural language processing and text mining of symptoms from electronic patient-authored text data', *International journal of medical informatics* **125**, 37–46.

Duncan, O. D. and Duncan, B. (1955), 'A methodological analysis of segregation indexes', *American sociological review* **20**(2), 210–217.

El Ouirdi, M., Segers, J., El Ouirdi, A. and Pais, I. (2015), 'Predictors of job seekers' self-disclosure on social media', *Computers in Human Behavior* **53**, 1–12.

Eriksson, S. and Lagerström, J. (2012), 'The labor market consequences of gender differences in job search', *Journal of Labor Research* **33**(3), 303–327.

Exley, C. L., Niederle, M. and Vesterlund, L. (2020), 'Knowing when to ask: The cost of leaning in', *Journal of Political Economy* **128**(3), 816–854.

Faberman, R. J. and Kudlyak, M. (2019), 'The intensity of job search and search duration', *American Economic Journal: Macroeconomics* **11**(3), 327–57.

Fishman, M. J. and Hagerty, K. M. (1989), 'Disclosure decisions by firms and the competition for price efficiency', *The Journal of Finance* **44**(3), 633–646.

Fishman, M. J. and Hagerty, K. M. (2003), 'Mandatory versus voluntary disclosure in markets with informed and uninformed customers', *Journal of Law, Economics, and organization* **19**(1), 45–63.

Fitzpatrick, M. A., Mulac, A. and Dindia, K. (1995), ‘Gender-preferential language use in spouse and stranger interaction’, *Journal of Language and Social Psychology* **14**(1-2), 18–39.

Fix, M., Galster, G. C. and Struyk, R. J. (1993), ‘An overview of auditing for discrimination’, *Clear and convincing evidence: Measurement of discrimination in America* **1**, 13–14.

Flory, J. A., Leibbrandt, A. and List, J. A. (2015), ‘Do competitive workplaces deter female workers? a large-scale natural field experiment on job entry decisions’, *The Review of Economic Studies* **82**(1), 122–155.

Fu, Z., Xian, Y., Gao, R., Zhao, J., Huang, Q., Ge, Y., Xu, S., Geng, S., Shah, C., Zhang, Y. et al. (2020), Fairness-aware explainable recommendation over knowledge graphs, *in* ‘Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval’, pp. 69–78.

Fuster, A., Goldsmith-Pinkham, P., Ramadorai, T. and Walther, A. (2020), ‘Predictably unequal? the effects of machine learning on credit markets’, *The Effects of Machine Learning on Credit Markets (October 1, 2020)* .

Gaddis, S. M. (2018), An introduction to audit studies in the social sciences, *in* ‘Audit studies: Behind the scenes with theory, method, and nuance’, Springer, pp. 3–44.

Gao, R., Ge, Y. and Shah, C. (2021), ‘Fair: Fairness-aware information retrieval evaluation’, *arXiv preprint arXiv:2106.08527* .

Gastil, J. (1990), ‘Generic pronouns and sexist language: The oxymoronic character of masculine generics’, *Sex roles* **23**(11), 629–643.

Gaucher, D., Friesen, J. and Kay, A. C. (2011), ‘Evidence that gendered wording in job advertisements exists and sustains gender inequality.’, *Journal of personality and social psychology* **101**(1), 109.

Ge, Y., Liu, S., Gao, R., Xian, Y., Li, Y., Zhao, X., Pei, C., Sun, F., Ge, J., Ou, W. et al. (2021), Towards long-term fairness in recommendation, *in* ‘Proceedings of the 14th ACM International Conference on Web Search and Data Mining’, pp. 445–453.

Gee, L. K. (2019), ‘The more you know: Information effects on job application rates in a large field experiment’, *Management Science* **65**(5), 2077–2094.

Gibbons, R. and Murphy, K. J. (1992), ‘Optimal incentive contracts in the presence of career concerns: Theory and evidence’, *Journal of political Economy* **100**(3), 468–505.

Goldin, C. and Rouse, C. (2000), ‘Orchestrating impartiality: The impact of" blind" auditions on female musicians’, *American economic review* **90**(4), 715–741.

Graham, P. (2002), ‘A plan for spam (http://www. paulgraham. com/spam. html)’.

Grossman, S. J. (1981), 'The informational role of warranties and private disclosure about product quality', *The Journal of Law and Economics* **24**(3), 461–483.

Guo, L. and Zhao, Y. (2009), 'Voluntary quality disclosure and market interaction', *Marketing Science* **28**(3), 488–501.

Hall, R. E. and Krueger, A. B. (2012), 'Evidence on the incidence of wage posting, wage bargaining, and on-the-job search', *American Economic Journal: Macroeconomics* **4**(4), 56–67.

Hannák, A., Wagner, C., Garcia, D., Mislove, A., Strohmaier, M. and Wilson, C. (2017), Bias in online freelance marketplaces: Evidence from taskrabbit and fiverr, *in* 'Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing', pp. 1914–1933.

Hansen, B. and McNichols, D. (2020), Information and the persistence of the gender wage gap: Early evidence from california's salary history ban, Technical report, National Bureau of Economic Research.

Hayashi, R. and Matsuda, N. (2020), 'Covid-19 impact on job postings: Real-time assessment using bangladesh and sri lanka online job portals'.

He, H., Neumark, D. and Weng, Q. (2021), 'Do workers value flexible jobs? a field experiment', *Journal of Labor Economics* **39**(3), 709–738.

Helleseter, M. D., Kuhn, P. and Shen, K. (2020), 'The age twist in employers' gender requests evidence from four job boards', *Journal of Human Resources* **55**(2), 428–469.

Hershbein, B. and Kahn, L. B. (2018), 'Do recessions accelerate routine-biased technological change? evidence from vacancy postings', *American Economic Review* **108**(7), 1737–72.

Hirshleifer, D. A., Lim, S. S. and Teoh, S. H. (2004), 'Disclosure to an audience with limited attention', *Available at SSRN 604142* .

Hoffman, C. and Hurst, N. (1990), 'Gender stereotypes: Perception or rationalization?', *Journal of personality and social psychology* **58**(2), 197.

Hong, W., Zheng, S., Wang, H. and Shi, J. (2013), 'A job recommender system based on user clustering.', *J. Comput.* **8**(8), 1960–1967.

Horton, J. J. (2017), 'The effects of algorithmic labor market recommendations: Evidence from a field experiment', *Journal of Labor Economics* **35**(2), 345–385.

Hotz, V. J. and Xiao, M. (2013), 'Strategic information disclosure: The case of multiattribute products with heterogeneous consumers', *Economic Inquiry* **51**(1), 865–881.

Ibañez, M. and Riener, G. (2018), 'Sorting through affirmative action: three field experiments in colombia', *Journal of Labor Economics* **36**(2), 437–478.

Jackson, N. and Lilleker, D. (2011), 'Microblogging, constituency service and impression management: Uk mps and the use of twitter', *The journal of legislative studies* **17**(1), 86–105.

Jannach, D., Resnick, P., Tuzhilin, A. and Zanker, M. (2016), 'Recommender systems—beyond matrix completion', *Communications of the ACM* **59**(11), 94–102.

Jannach, D., Zanker, M., Felfernig, A. and Friedrich, G. (2010), *Recommender systems: an introduction*, Cambridge University Press.

Jiang, R., Chiappa, S., Lattimore, T., György, A. and Kohli, P. (2019), Degenerate feedback loops in recommender systems, *in* 'Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society', pp. 383–390.

Jiang, S., Hu, Y., Kang, C., Daly Jr, T., Yin, D., Chang, Y. and Zhai, C. (2016), Learning query and document relevance from a web-scale click graph, *in* 'Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval', pp. 185–194.

Jin, G. Z. (2005), 'Competition and disclosure incentives: an empirical study of hmos', *Rand journal of Economics* pp. 93–112.

Jin, G. Z. and Leslie, P. (2003), 'The effect of information on product quality: Evidence from restaurant hygiene grade cards', *The Quarterly Journal of Economics* **118**(2), 409–451.

Jovanovic, B. (1982), 'Truthful disclosure of information', *The Bell Journal of Economics* pp. 36–44.

Kay, M., Matuszek, C. and Munson, S. A. (2015), Unequal representation and gender stereotypes in image search results for occupations, *in* 'Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems', pp. 3819–3828.

Keith, K. and McWilliams, A. (1999), 'The returns to mobility and job search by gender', *ILR Review* **52**(3), 460–477.

Kellogg, K. C., Valentine, M. A. and Christin, A. (2020), 'Algorithms at work: The new contested terrain of control', *Academy of Management Annals* **14**(1), 366–410.

Khanna, S. (2020), 'Salary history bans and wage bargaining: Experimental evidence', *Labour Economics* p. 101853.

Kilby, A. E. (2021), Algorithmic fairness in predicting opioid use disorder using machine learning, *in* 'Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency', pp. 272–272.

Kim, M. (2015), 'Pay secrecy and the gender wage gap in the united states', *Industrial Relations: A Journal of Economy and Society* **54**(4), 648–667.

Kim, P. T. (2017), 'Auditing algorithms for discrimination', *U. Pa. L. Rev. Online* **166**, 189.

Klare, B. F., Burge, M. J., Klontz, J. C., Bruegge, R. W. V. and Jain, A. K. (2012), 'Face recognition performance: Role of demographic information', *IEEE Transactions on Information Forensics and Security* **7**(6), 1789–1801.

Kleinberg, J., Ludwig, J., Mullainathan, S. and Rambachan, A. (2018), Algorithmic fairness, *in* 'Aea papers and proceedings', Vol. 108, pp. 22–27.

Köchling, A. and Wehner, M. C. (2020), 'Discriminated by an algorithm: a systematic review of discrimination and fairness by algorithmic decision-making in the context of hr recruitment and hr development', *Business Research* pp. 1–54.

Krause, A., Rinne, U. and Zimmermann, K. F. (2012), 'Anonymous job applications in europe', *IZA Journal of European Labor Studies* **1**(1), 1–20.

Kreisman, D., Smith, J., Arifin, B. et al. (2021), Labor market signaling and the value of college: Evidence from resumes and the truth, Technical report, Institute of Labor Economics (IZA).

Kroft, K., Lange, F. and Notowidigdo, M. J. (2013), 'Duration dependence and labor market conditions: Evidence from a field experiment', *The Quarterly Journal of Economics* **128**(3), 1123–1167.

Kudlyak, M., Lkhagvasuren, D. and Sysuyev, R. (2013), 'Systematic job search: New evidence from individual job application data'.

Kuhn, P. (1992), 'Mandatory notice', *Journal of Labor Economics* **10**(2), 117–137.

Kuhn, P. J. and Shen, K. (2021), What happens when employers can no longer discriminate in job ads?, Technical report, National Bureau of Economic Research.

Kuhn, P. and Shen, K. (2013), 'Gender discrimination in job ads: Evidence from china', *The Quarterly Journal of Economics* **128**(1), 287–336.

Kuhn, P., Shen, K. and Zhang, S. (2020), 'Gender-targeted job ads in the recruitment process: Facts from a chinese job board', *Journal of Development Economics* **147**, 102531.

Lambrecht, A. and Tucker, C. (2019), 'Algorithmic bias? an empirical study of apparent gender-based discrimination in the display of stem career ads', *Management science* **65**(7), 2966–2981.

Lang, K., Manove, M. and Dickens, W. T. (2005), 'Racial discrimination in labor markets with posted wage offers', *American Economic Review* **95**(4), 1327–1340.

Lazear, E. P., Shaw, K. L. and Stanton, C. (2016), 'Making do with less: working harder during recessions', *Journal of Labor Economics* **34**(S1), S333–S360.

Lazear, E. P., Shaw, K. L. and Stanton, C. T. (2018), 'Who gets hired? the importance of competition among applicants', *Journal of Labor Economics* **36**(S1), S133–S181.

Lee, D. H. and Brusilovsky, P. (2007), Fighting information overflow with personalized comprehensive information access: A proactive job recommender, *in* 'Third International Conference on Autonomic and Autonomous Systems (ICAS'07)', IEEE, pp. 21–21.

Lee, M. K. and Baykal, S. (2017), Algorithmic mediation in group decisions: Fairness perceptions of algorithmically mediated vs. discussion-based social division, *in* 'Proceedings of the 2017 acm conference on computer supported cooperative work and social computing', pp. 1035–1048.

Leibbrandt, A. and List, J. A. (2015), 'Do women avoid salary negotiations? evidence from a large-scale natural field experiment', *Management Science* **61**(9), 2016–2024.

Leibbrandt, A. and List, J. A. (2018), Do equal employment opportunity statements backfire? evidence from a natural field experiment on job-entry decisions, Technical report, National Bureau of Economic Research.

Levin, D., Peck, J. and Ye, L. (2009), 'Quality disclosure and competition', *The Journal of Industrial Economics* **57**(1), 167–196.

Li, D., Raymond, L. R. and Bergman, P. (2020), Hiring as exploration, Technical report, National Bureau of Economic Research.

Lindqvist, A., Renström, E. A. and Sendén, M. G. (2019), 'Reducing a male bias in language? establishing the efficiency of three different gender-fair language strategies', *Sex Roles* **81**(1), 109–117.

Logan, J. R., Stults, B. J. and Farley, R. (2004), 'Segregation of minorities in the metropolis: Two decades of change', *Demography* **41**(1), 1–22.

Mann, G. and O'Neil, C. (2016), 'Hiring algorithms are not neutral', *Harvard Business Review* **9**, 2016.

Manning, A. (2003), 'Monopsony in motion: imperfect competition in labor markets princeton university press', *Princeton NJ* .

Marinescu, I. and Rathelot, R. (2018*a*), 'Mismatch unemployment and the geography of job search', *American Economic Journal: Macroeconomics* **10**(3), 42–70.

Marinescu, I. and Rathelot, R. (2018*b*), 'Mismatch unemployment and the geography of job search', *American Economic Journal: Macroeconomics* **10**(3), 42–70.
**URL:** *https://www.aeaweb.org/articles?id=10.1257/mac.20160312*

Marinescu, I. and Wolthoff, R. (2020), 'Opening the black box of the matching function: The power of words', *Journal of Labor Economics* **38**(2), 535–568.

Mas, A. and Pallais, A. (2017), 'Valuing alternative work arrangements', *American Economic Review* **107**(12), 3722–59.

Mathios, A. D. (2000), 'The impact of mandatory disclosure laws on product choices: An analysis of the salad dressing market', *The Journal of Law and Economics* **43**(2), 651–678.

Matsuda, N., Ahmed, T. and Nomura, S. (2019), 'Labor market analysis using big data: The case of a pakistani online job portal', *World Bank Policy Research Working Paper* (9063).

McCallum, A., Nigam, K. et al. (1998), A comparison of event models for naive bayes text classification, *in* 'AAAI-98 workshop on learning for text categorization', Vol. 752, Citeseer, pp. 41–48.

Milgrom, P. R. (1981), 'Good news and bad news: Representation theorems and applications', *The Bell Journal of Economics* pp. 380–391.

Mitchell, T. M. (1997), 'Does machine learning really work?', *AI magazine* **18**(3), 11–11.

Modestino, A. S., Shoag, D. and Ballance, J. (2016), 'Downskilling: changes in employer skill requirements over the business cycle', *Labour Economics* **41**, 333–347.

Mullainathan, S. and Obermeyer, Z. (2021), On the inequity of predicting a while hoping for b, *in* 'AEA Papers and Proceedings', Vol. 111, pp. 37–42.

Neumark, D., Burn, I. and Button, P. (2019), 'Is it harder for older workers to find jobs? new and improved evidence from a field experiment', *Journal of Political Economy* **127**(2), 922–970.

Niederle, M. and Vesterlund, L. (2007), 'Do women shy away from competition? do men compete too much?', *The quarterly journal of economics* **122**(3), 1067–1101.

Ningrum, P. K., Pansombut, T. and Ueranantasun, A. (2020), 'Text mining of online job advertisements to identify direct discrimination during job hunting process: A case study in indonesia', *Plos one* **15**(6), e0233746.

Nomura, S., Imaizumi, S., Areias, A. C. and Yamauchi, F. (2017), 'Toward labor market policy 2.0: the potential for using online job-portal big data to inform labor market policies in india', *World Bank Policy Research Working Paper* (7966).

Nosek, B. A., Banaji, M. R. and Greenwald, A. G. (2002), 'Harvesting implicit group attitudes and beliefs from a demonstration web site.', *Group Dynamics: Theory, Research, and Practice* **6**(1), 101.

Obermeyer, Z., Powers, B., Vogeli, C. and Mullainathan, S. (2019), 'Dissecting racial bias in an algorithm used to manage the health of populations', *Science* **366**(6464), 447–453.

OECD (2016), *The survey of adult skills: Reader's companion*, Organisation for Economic Co-operation and Development.
**URL:** *https://doi.org/10.1787/9789264204027-en*

Oyer, P. and Schaefer, S. (2010), Personnel economics: Hiring and incentives, Technical report, National Bureau of Economic Research.

Parsons, D. O. (1991), 'The job search behavior of employed youth', *The Review of Economics and Statistics* pp. 597–604.

PayScale (2017), '*Is Asking for Salary History ...History?*'. https://www.payscale.com/data/salary-history.

Pedreshi, D., Ruggieri, S. and Turini, F. (2008), Discrimination-aware data mining, *in* 'Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining', pp. 560–568.

Pekkarinen, T. and Vartiainen, J. (2006), 'Gender differences in promotion on a job ladder: Evidence from finnish metalworkers', *ILR Review* **59**(2), 285–301.

Petit, P. (2007), 'The effects of age and family constraints on gender hiring discrimination: A field experiment in the french financial sector', *Labour Economics* **14**(3), 371–391.

Pető, R. and Reizer, B. (2021), 'Gender differences in the skill content of jobs', *Journal of Population Economics* **34**(3), 825–864.

Poutanen, S. and Kovalainen, A. (2017), New economy, platform economy and gender, *in* 'Gender and Innovation in the New Economy', Springer, pp. 47–96.

Powell Jr, L. F. (1972), 'Pittsburgh press co. v. pittsburgh commission on human relations, et. al.,u.s. supreme court decision argued march 20. 1973, decided on june 21, 1973.'.

Prassl, J. (2018), *Humans as a service: The promise and perils of work in the gig economy*, Oxford University Press.

Raghavan, M., Barocas, S., Kleinberg, J. and Levy, K. (2020), Mitigating bias in algorithmic hiring: Evaluating claims and practices, *in* 'Proceedings of the 2020 conference on fairness, accountability, and transparency', pp. 469–481.

Rambachan, A., Kleinberg, J., Ludwig, J. and Mullainathan, S. (2020), An economic perspective on algorithmic fairness, *in* 'AEA Papers and Proceedings', Vol. 110, pp. 91–95.

Rambachan, A. and Roth, J. (2019), 'Bias in, bias out? evaluating the folk wisdom', *arXiv preprint arXiv:1909.08518* .

Riach, P. A. and Rich, J. (2006), 'An experimental investigation of sexual discrimination in hiring in the english labor market', *Advances in Economic Analysis & Policy* **5**(2).

Rich, J. (2014), 'What do field experiments of discrimination in markets tell us? a meta analysis of studies conducted since 2000'.

Richardson, M., Dominowska, E. and Ragno, R. (2007), Predicting clicks: estimating the click-through rate for new ads, *in* 'Proceedings of the 16th international conference on World Wide Web', pp. 521–530.

Roberts, D. C. and Utych, S. M. (2020), 'Linking gender, language, and partisanship: Developing a database of masculine and feminine words', *Political Research Quarterly* **73**(1), 40–50.

Roussille, N. (2020), 'The central role of the ask gap in gender pay inequality', *URL: https://ninaroussille. github. io/files/Roussille_ askgap. pdf* .

Rudman, L. A. and Kilianski, S. E. (2000), 'Implicit and explicit attitudes toward female authority', *Personality and social psychology bulletin* **26**(11), 1315–1328.

Sandvig, C., Hamilton, K., Karahalios, K. and Langbort, C. (2014), 'An algorithm audit', *Data and Discrimination: Collected Essays. Washington, DC: New America Foundation* pp. 6–10.

Savage, D. D. and Bales, R. (2016), 'Video games in job interviews: using algorithms to minimize discrimination and unconscious bias', *ABAJ Lab. & Emp. L.* **32**, 211.

Shen, K. and Kuhn, P. (2013), Do chinese employers avoid hiring overqualified workers? evidence from an internet job board, *in* 'Labor Market Issues in China', Emerald Group Publishing Limited.

Siting, Z., Wenxing, H., Ning, Z. and Fan, Y. (2012), Job recommender systems: a survey, *in* '2012 7th International Conference on Computer Science & Education (ICCSE)', IEEE, pp. 920–924.

Sonboli, N., Smith, J. J., Cabral Berenfus, F., Burke, R. and Fiesler, C. (2021), Fairness and transparency in recommendation: The users' perspective, *in* 'Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization', pp. 274–279.

Sorensen, E. (1990), 'The crowding hypothesis and comparable worth', *Journal of Human Resources* pp. 55–89.

Stinebrickner, T. R., Stinebrickner, R. and Sullivan, P. J. (2018), Job tasks and the gender wage gap among college graduates, Technical report, National Bureau of Economic Research.

Stuhlmacher, A. F. and Walters, A. E. (1999), 'Gender differences in negotiation outcome: A meta-analysis', *Personnel Psychology* **52**(3), 653–677.

Tambe, P., Cappelli, P. and Yakubovich, V. (2019), 'Artificial intelligence in human resources management: Challenges and a path forward', *California Management Review* **61**(4), 15–42.

Topel, R. H. and Ward, M. P. (1992), 'Job mobility and the careers of young men', *The Quarterly Journal of Economics* **107**(2), 439–479.

Waldman, M. (2013), 'Classic promotion tournaments versus market-based tournaments', *International Journal of Industrial Organization* **31**(3), 198–210.

Walsh, J. et al. (1975), 'Help wanted: Case studies of classified ads.'.

Watch/Asia, H. R. (2018), *" Only Men Need Apply": Gender Discrimination in Job Advertisements in China*, Human Rights Watch.

Yu, H., Liu, C. and Zhang, F. (2011), 'Reciprocal recommendation algorithm for the field of recruitment', *Journal of Information & Computational Science* **8**(16), 4061–4068.

Zhou, X., Zhang, J. and Song, X. (2013), 'Gender discrimination in hiring: Evidence from 19,130 resumes in china', *Available at SSRN 2195840* .

Zide, J., Elman, B. and Shahani-Denning, C. (2014), 'Linkedin and recruitment: How profiles differ across occupations', *Employee Relations* **36**(5), 583–604.

Zoloth, B. S. (1976), 'Alternative measures of school segregation', *Land Economics* **52**(3), 278–298.